Volume 63

Advances in Child Development and Behavior

SERIES EDITOR Jeffrey J. Lockman

VOLUME SIXTY THREE

ADVANCES IN CHILD DEVELOPMENT AND BEHAVIOR

ADVANCES IN CHILD DEVELOPMENT AND BEHAVIOR

Series Editor

JEFFREY J. LOCKMAN
Department of Psychology,
Tulane University,
New Orleans, LA, United States

VOLUME SIXTY THREE

ADVANCES IN CHILD DEVELOPMENT AND BEHAVIOR

Edited by

JEFFREY J. LOCKMAN
Department of Psychology,
Tulane University,
New Orleans, LA, United States

ACADEMIC PRESS
An imprint of Elsevier
ELSEVIER

Academic Press is an imprint of Elsevier
50 Hampshire Street, 5th Floor, Cambridge, MA 02139, United States
525 B Street, Suite 1650, San Diego, CA 92101, United States
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, United Kingdom
125 London Wall, London EC2Y 5AS, United Kingdom

First edition 2022

ISBN: 978-0-323-99076-9
ISSN: 0065-2407 (Series)

For information on all Academic Press publications
visit our website at https://www.elsevier.com/books-and-journals

Publisher: Zoe Kruze
Acquisitions Editor: Mariana Kuhl
Developmental Editor: Naiza Ermin Mendoza
Production Project Manager: Abdulla Sait
Cover Designer: Miles Hitchen

Typeset by STRAIVE, India

Contents

Contributors

Kim Astor
Uppsala Child & Baby Lab, Uppsala University, Uppsala, Sweden

Julie C. Bowker
University at Buffalo, The State University of New York, Buffalo, NY, United States

Gustavo Carlo
School of Education, University of California, Irvine, CA, United States

Eveline A. Crone
Erasmus School of Social and Behavioral Sciences, Erasmus University Rotterdam, Rotterdam; Department of Developmental and Educational Psychology, Leiden University, Leiden, The Netherlands

Alexandra N. Davis
University of New Mexico, Albuquerque, NM, United States

Gustaf Gredebäck
Uppsala Child & Baby Lab, Uppsala University, Uppsala, Sweden

Megan R. Gunnar
Institute of Child Development, University of Minnesota-Twin Cities, Minneapolis, MN, United States

Dale F. Hay
Cardiff University, Cardiff, United Kingdom

Mariann A. Howland
Institute of Child Development, University of Minnesota-Twin Cities, Minneapolis, MN, United States

George P. Knight
Arizona State University, Tempe, AZ, United States

Rongzhi Liu
Department of Psychology, University of California, Berkeley, CA, United States

Amy Paine
Cardiff University, Cardiff, United Kingdom

Charlotte J. Patterson
University of Virginia, Charlottesville, VA, United States

Tobias Richter
Department of Psychology, University of Würzburg, Würzburg, Germany

Charlotte Robinson
Cardiff University, Cardiff, United Kingdom

Meredith L. Rowe
Harvard Graduate School of Education, Harvard University, Cambridge, MA, United States

Wolfgang Schneider
Department of Psychology, University of Würzburg, Würzburg, Germany

Robert S. Siegler
Teachers College, Columbia University, New York, NY, United States

Sophie W. Sweijen
Erasmus School of Social and Behavioral Sciences, Erasmus University Rotterdam, Rotterdam, The Netherlands

Lysanne W. te Brinke
Erasmus School of Social and Behavioral Sciences, Erasmus University Rotterdam, Rotterdam, The Netherlands

Jing Tian
Temple University, Philadelphia, PA, United States

Catharina Tibken
Department of Psychology, University of Würzburg, Würzburg, Germany

Suzanne van de Groep
Erasmus School of Social and Behavioral Sciences, Erasmus University Rotterdam, Rotterdam, The Netherlands

Jenna Weingarten
University at Buffalo, The State University of New York, Buffalo, NY, United States

Fei Xu
Department of Psychology, University of California, Berkeley, CA, United States

Preface

Volume 63 of *Advances in Child Development and Behavior* continues the rich tradition established in 1963 by Lew Lipsitt and Charles Spiker, founding editors of the *Advances* series. In their preface (1963, p. vii) to Volume 1 of this series, they wrote:

> *The serial publication of Advances in Child Development and Behavior is intended to provide scholarly reference articles in the field and to serve two purposes. On the one hand, it is hoped that teachers, research workers, and students will find these critical syntheses useful in the endless task of keeping abreast of growing knowledge in areas peripheral to their primary focus of interest. There is currently an indisputable need for technical, documented reviews which would facilitate this task by reducing the frequency with which original papers must be consulted, particularly in such secondary areas. On the other hand, the editors are also convinced that research in child development has progressed to the point that such integrative and critical papers will be of considerable usefulness to researchers within problem areas of great concern to their own research programs.*

Similar to Volume 1 of the *Advances* series, Volume 63 features overviews and critical analyses of research areas in developmental science. These reviews are directed to researchers, educators, policy-makers, and students, all of whom are either specialists in a research area and/or share an abiding interest in the science of child development. Given what often seems like the exponential growth of research in developmental science, I am convinced that the *Advances* series is needed more than ever to help synthesize current knowledge and build on the resulting foundation to ensure the future growth of our field.

Volume 63 comprises 11 chapters, which together span the infancy to adolescence period. The chapters highlight theoretical and substantive advances in research that are diverse with respect to populations and methods. Notably, each chapter considers the reciprocal relation between research and translation.

In Chapter 1, Siegler and Tian address children's understanding of percentages within the context of overall numerical development. By integrating theory with educational practice, Siegler and Tian's programmatic studies illuminate why children often encounter difficulties in solving percentage problems. Their findings suggest straightforward ways to promote children's understanding of percentages through appropriately designed instruction and materials. In Chapter 2, Gunnar and Howland provide a truly developmental overview of recent work on calibration and recalibration of stress response systems in humans. Their masterful synthesis offers new insights into the possibility of recalibration of some of these systems at different points during the life span, thereby suggesting new strategies for intervention. In Chapter 3, Patterson assesses what has been learned about lesbian, gay, bisexual, transgender, and queer (LGBTQ+) parents and the development of their children, especially in light of the dramatic increase in the number of these families around the world. Despite the obstacles faced by many LGBTQ+ families because of societal laws and attitudes, Patterson's review demonstrates that these families are resilient and that children from these families typically thrive. By providing a roadmap for future research on LGBTQ+ families, Patterson shows how such research can broaden the understanding of parenting and child development.

Chapter 4, by Rowe, is also situated at the intersection of research and policy. Rowe asks how research on the role of experience in language and literacy development can be leveraged to reduce socioeconomically associated differences in these skills, which are already evident when children first enter school. Rowe suggests that policies that are aimed at enhancing high school students' knowledge about child development coupled with policies that reduce parental stress can lead to improved quality and quantity of language input to children and thus boost child language and academic outcomes.

In the next two chapters, prosocial development during the adolescent period is considered. In Chapter 5, Carlo and Knight highlight a fundamental challenge faced by many Latinx youth in the United States: daily and long-term exposure to social and structural inequities. Exposure to these inequities increases developmental risk and reduces developmental opportunity. To help combat the effects of these inequities, Carlo and Knight advocate for a developmental strengths approach that promotes prosocial behaviors between diverse youth, especially the majority and minority members of our society. In Chapter 6, Crone, Sweijen, te Brinke, and van de Groep also focus on prosocial behaviors in adolescence. They offer a comprehensive review of the behavioral and neural pathways that underlie the development of prosocial behaviors during adolescence by focusing on affective and sociocognitive domains. Their review has important implications for the design of effective interventions that foster prosocial behaviors and, more broadly, resilience and adaption throughout the life span.

In the next three chapters, common assumptions regarding the development of joint attention (Chapter 7) and peer interaction (Chapters 8 and 9) are reexamined. In Chapter 7, Astor and Gredebäck review work on the early development of gaze following. As Astor and Gredebäck note, research on gaze following in infancy has been plentiful. They argue, however, that more work needs to be focused on the mechanisms that underlie the development of gaze following during early development. By identifying the unresolved issues in the gaze following literature, Astor and Gredebäck's essay can advance research in this area. In Chapter 8, Hay, Paine, and Robinson take up the relation between cooperation and conflict in very young children. Hay et al. observe that although these two forms of social interaction are often studied in isolation from one another, the two forms of interaction often occur in concert and constitute part of the ongoing dynamic of young children's social exchanges. As the authors demonstrate through their longitudinal research, social interchanges populated by cooperation and conflict offer rich opportunities to refine social, cognitive, and communication skills during the early childhood years. In Chapter 9, Bowker and Weingarten also argue for a new wave of research on the development of children's friendships. In contrast to a good deal of research on children's friendships that examines these relationships at one point in time, they advocate for an approach that emphasizes temporal changes over developmental time. Such an approach can generate new insights into the processes of friendship formation, dissolution, and re-formation during different periods of development and inform intervention efforts.

The final two chapters address ways in which children understand their own thinking abilities and how they reason about other people. In Chapter 10, Schneider, Tibken, and Richter discuss the development of metacognition. In their systematic review, they consider both the declarative and procedural components of this ability. Although work on metacognition originated in the field of memory development, Schneider et al. show how subsequent research on metacognition has been applied to different academic domains, including reading comprehension. Their work reveals how a focus on metacognitive skills in classroom settings can lead to improvement in children's school performance. Finally, in Chapter 11, Liu and Xu review developmental models that address how very young children reason about others. Although the literature is replete with findings on this topic, Liu and Xu observe that the literature is nevertheless short on research that directly addresses the mechanisms that drive developmental growth in this area. To fill this important gap, Liu and Xu discuss how Bayesian probabilistic models may be used to deepen our understanding of these mechanisms. Liu and Xu's integrative approach paves the way for future work that can illuminate how young children's reasoning about others' mental states and their actions develops.

In conclusion, this volume represents not only the contributions of the authors whose chapters appear before you, but also the work of many others behind the scenes at Elsevier. Chief among them is Naiza Mendoza, Developmental Editor at Elsevier, without whose efforts this volume would not have been possible. My deepest thanks to Naiza for her dedication in helping bring this volume to fruition.

Finally, at the end of September 2021, Lewis P. Lipsitt, one of the founding editors of this series, and an overall visionary in the field of developmental science, passed away. Lew was a prolific, careful, and generous scholar; he shaped many fields of inquiry in developmental science, especially the field of infancy as we currently know it. His far-ranging contributions live on as well in this series. It is to his work and memory that I dedicate Volume 63 of *Advances in Child Development and Behavior*.

Jeffrey J. Lockman
Department of Psychology
Tulane University
New Orleans, LA, United States

CHAPTER ONE

Why do we have three rational number notations? The importance of percentages

Robert S. Siegler[a,*] **and Jing Tian**[b]
[a]Teachers College, Columbia University, New York, NY, United States
[b]Temple University, Philadelphia, PA, United States
[*]Corresponding author: e-mail address: rss2169@tc.columbia.edu

Contents

Abstract

The integrated theory of numerical development provides a unified approach to understanding numerical development, including acquisition of knowledge about whole numbers, fractions, decimals, percentages, negatives, and relations among all of these types of numbers (Siegler, Thompson, & Schneider, 2011). Although, considerable progress has been made toward many aspects of this integration (Siegler, Im, Schiller, Tian, & Braithwaite, 2020), the role of percentages has received much less attention than that of the other types of numbers. This chapter is an effort to redress this imbalance by reporting data on understanding of percentages and their relations to other types of numbers. We first describe the integrated theory; then summarize what is known about development of understanding of whole numbers, fractions, and decimals; then describe recent progress in understanding the role of percentages; and finally consider instructional implications of the theory and research.

Advances in Child Development and Behavior, Volume 63
ISSN 0065-2407
https://doi.org/10.1016/bs.acdb.2022.05.001

1. The integrated theory of numerical development

Most prominent theories of numerical development focus on knowledge of whole numbers. This is true of privileged domain theories (e.g., Spelke & Kinzler, 2007), conceptual change theories (e.g., Carey, 2009), evolutionary theories (e.g., Geary, 2006), and information processing theories (e.g., Fias, Sahan, Ansari, & Lyons, 2021). Rational numbers have received far less emphasis, and usually have been used primarily as contrastive cases when they have received attention. For example, within privileged domains, evolutionary, and conceptual change theories, the rapid, effortless, and universal development of basic understanding of whole numbers is often contrasted to the gradual, painstaking, and far-from-universal grasp of rational numbers (e.g., Geary, 2004; Gelman & Williams, 1998; Wynn, 2002).

Although differences between acquisition of understanding of whole and rational numbers are important, they are only part of the story of numerical development. Along with the differences, there are strong commonalities in development across different types of numbers. The integrated theory focuses on both the commonalities and the differences in acquisition of understanding of different types of numbers. It also focuses on commonalities as well as differences in development of understanding of individual numbers and arithmetic combinations of numbers of each type.

The basic tenet of the integrated theory is that numerical development is fundamentally a process of acquiring increasingly precise knowledge of the magnitudes of an increasingly broad range of numbers (Siegler et al., 2011). That is, numerical development involves learning the relations of numbers to their magnitudes. This process occurs earlier with non-symbolic than with symbolic numbers, with smaller symbolic whole numbers than with larger symbolic whole numbers, and with whole numbers than with rational numbers. Fig. 1 illustrates the hypothesized developmental progression.

The importance of magnitude knowledge extends to arithmetic combinations of numbers as well as to individual numbers. Consider potential reactions to a common fraction arithmetic error, separately adding numerators and denominators (e.g., claiming that $2/3 + 1/7 = 3/10$.) A child who understood the magnitudes of these fractions and who knew that sums of positive numbers must be greater than the individual addends would know that this answer was impossible, because the operand 2/3 is greater than the proposed sum 3/10. Given that many children know and use both this incorrect fraction addition strategy and the correct approach (Siegler & Pyke, 2013), attention to magnitudes could increasingly lead children to

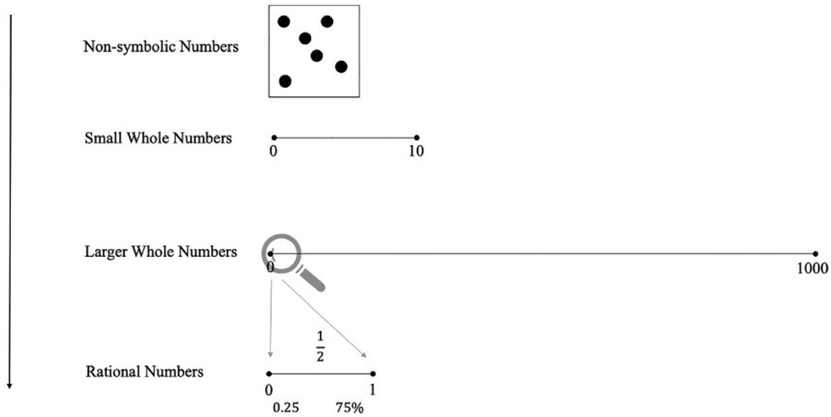


Fig. 1 Numerical development progression as proposed by the integrated theory of numerical development.

reject the incorrect strategy and rely on the correct one. In contrast, children who did not know or ignored the magnitudes of the operands and sum might be satisfied with this answer, because "2 + 1" does equal "3" and "3 + 7" does equal "10."

The process of learning about magnitudes generally occurs earlier, and to a higher asymptotic level, with the magnitudes of individual numbers than with the magnitudes of arithmetic combinations of numbers. This is true with both whole and rational numbers, though the degree of difference is much greater with fractions (and perhaps other rational numbers) than with whole numbers. Evidence for this claim was provided by Braithwaite, Tian, and Siegler (2018) who found that number line estimates of 6th and 7th graders for individual whole numbers were somewhat more accurate than estimates for individual fractions, whereas estimates for whole number sums were far more accurate than those for fraction sums.

In addition to understanding relations between various types of numbers and their magnitudes, children also need to learn which properties of whole numbers extend to other types of numbers and which do not. One similarity that children need to understand is that adding all types of positive numbers, not just whole numbers, always results in a sum larger than any of the addends. One difference that they need to learn is that multiplying positive rational numbers, unlike multiplying whole numbers, does not always result in products larger than the multiplicands. In particular, multiplying numbers between zero and one never results in a product larger than any of the multiplicands.

A considerable body of evidence consistent with the integrated theory of numerical development has emerged. Accuracy of number line estimates of the magnitudes of individual numbers is strongly correlated with accuracy of solutions to arithmetic problems for both whole numbers (Castronovo & Göbel, 2012; Gunderson, Ramirez, Beilock, & Levine, 2012; Linsen, Verschaffel, Reynvoet, & De Smedt, 2015) and fractions (Siegler et al., 2011; Torbeyns, Schneider, Xin, & Siegler, 2015). On magnitude comparison tasks, ratio dependence is present with fractions and decimals, just as it is with whole numbers (Hurst & Cordes, 2018). With regard to individual differences, children's knowledge of the magnitudes of whole numbers in first grade predicts their knowledge of the magnitudes of fractions in eighth grade, even after statistically controlling for the students' IQ, executive functioning, race and gender, as well as their parents' education and income (Bailey, Siegler, & Geary, 2014). Perhaps most compelling, interventions that improve children's knowledge of the magnitudes of individual whole numbers improve their knowledge of the magnitudes of whole number sums (Booth & Siegler, 2008; Siegler & Ramani, 2009), and interventions that improve children's knowledge of the magnitudes of individual fractions improve their knowledge of the magnitudes of fraction sums (Braithwaite & Siegler, 2021; Fazio, Kennedy, & Siegler, 2016). Thus, the integrated theory of numerical development is useful for characterizing the development of different types of numbers, both individual numbers and arithmetic combinations of numbers.

2. The importance of rational numbers

As noted previously, the integrated theory places greater emphasis on rational numbers than do alternative theories of numerical development. Both theoretical and practical considerations have led to this emphasis. A major theoretical reason is that learning rational numbers provides the first challenge to many children's assumption that properties of whole numbers are properties of all numbers (Gelman, 1991). This is a reasonable assumption when all or almost all of children's numerical experience has been with whole numbers, but many children (and even adults) continue to generalize properties of whole numbers to other types of numbers even after years of instruction demonstrating that some of the generalizations are wrong (Braithwaite & Siegler, 2018; Ni & Zhou, 2005). From the perspective of the integrated theory, comprehensive numerical understanding requires knowledge of which properties do and do not generalize and of why they do or do not.

Numerous applied considerations also argue for the integrated theory's emphasis on rational numbers. Such knowledge is vital to academic, occupational, and everyday competence. In academic contexts, knowledge of fractions and decimals is essential in math courses such as algebra, trigonometry, and statistics; science courses such as chemistry and physics; and social science courses, such as psychology, sociology, and economics. Consistent with this view, individual children's knowledge of fractions and decimals in elementary school is predictive of their later success in algebra and overall math achievement in high school, even after controlling for their whole number arithmetic knowledge, reading comprehension, IQ, working memory, and family background (e.g., Siegler et al., 2012).

Knowledge of rational numbers also is essential in occupational contexts. Among a nationally representative sample of more than 2300 US workers in a wide range of blue- and white-collar occupations, 68% reported using fractions in their jobs, vs 22% reporting use of any more advanced mathematics, such as algebra, geometry, or statistics (Handel, 2016). Weak understanding of decimals and fractions precludes many people from employment in well-paying occupations, such as pharmacist, nurse, and machinist (McCloskey, 2007; Sformo, 2008).

Understanding of rational numbers is also needed in a wide variety of everyday contexts. Among these contexts are adjusting recipes to feed a specific number of guests, adjusting doses of medications according to one's weight, and dividing pizzas and desserts into equal size portions.

Commensurate with their usefulness in academic, occupational, and everyday contexts, fractions and decimals receive prolonged attention in school. In the United States and many other countries, fractions are introduced in third or fourth grade; are a major topic of instruction in fourth, fifth, and sixth grades; and receive some attention in seventh grade. Decimals are introduced somewhat later and receive somewhat less instruction, but still a substantial amount. (Details of the textbook coverage are provided below.) Commensurate with this emphasis in school, many research studies of children's acquisition of knowledge about fractions and decimals have been conducted (Siegler et al., 2020).

Percentages are a different story.[a] Little is known about people's understanding of percentages or how that understanding develops. They receive far less coverage in textbooks, and the coverage they receive is provided later

[a] For present purposes we limit the term "percentages" to pure percentages, that is, ones without a decimal component; hybrids of percentages and decimals (e.g., 60.12%) are, of course, also possible.

than that of fractions and decimals. A recent review found fewer than 10 studies of children's knowledge of percentages (Tian & Siegler, 2018).

The paucity of knowledge about development of understanding of percentages is unfortunate. Informal observation suggests that percentages are used in many everyday situations, quite possibly more everyday contexts than fractions and decimals (Jacobs Danan & Gelman, 2017). Sales on consumer goods are typically described in percentage terms (e.g., a 20% off sale), as are classroom test scores (e.g., 85% correct), the remaining electric charge on phones and computers (5% left), and information in news stories (e.g., breakthrough cases of COVID-19 have occurred in 0.08% of people with two vaccinations). Percentages are also often used in arithmetic computations, such as calculating tips in restaurants and comparing prices (e.g., is a jacket a better buy after a 30% reduction from a price of $40 than after a 20% reduction from a price of $30).

Several interpretations of the discrepancy between the ubiquitous use of percentages and the paucity of their coverage in textbooks and research seem plausible. Perhaps, students become highly proficient with percentages after the limited coverage provided in textbooks. Perhaps, people prefer to think of proportional relations in terms of fractions and decimals rather than percentages. Perhaps, students are unable to learn about percentages regardless of the instruction they receive.

None these possibilities turn out to be true, though. As illustrated in the remainder of this chapter:

(1) Many children have only weak understanding of percentages.
(2) People often prefer to describe proportional relations as percentages rather than as fractions or decimals.
(3) US textbooks cover percentages far less than fractions or decimals and focus on different types of problems.
(4) Well-designed instruction that focuses on percentages can improve not only understanding of them but also of fractions and decimals.

3. Children's knowledge of percentages

Students' difficulty learning about percentages has been an enduring problem. Relevant data comes from a pair of studies conducted three-quarters of a century ago. Guiler (1946b) examined knowledge of percentage arithmetic of more than 900 ninth graders; Guiler (1946a) performed a parallel study with university students. All problems related three values: a whole (original value), a part (related value), and a percentage.

Note that in this usage, the part can be larger than the whole; for example, in the problem "What is 125% of 20," the whole is 20 and the part is 25. Thus, the whole is the base, and the part is defined relative to the whole. The part will be smaller than the whole if the percentage is below 100% (75% of 20) and larger than the whole if the percentage is above 100% (125% of 20).

As illustrated in this example, percentage arithmetic problems generally involve three variables, two of which are specified on each problem. On *percent-whole problems*, the percentage and whole are specified, and students need to find a part (e.g., 80% of \$2.00 = ____). In *whole-part problems*, the whole and the part are specified, and students need to find the percentage (e.g., 20 games is ____% of 25 games). In *percent-part problems*, the percentage and part are specified, and students need to find the whole (e.g., "125% of \$____ = \$8.00"). Guiler (1946b) presented the ninth graders two items of each type and scored their performance according to whether they "showed weaknesses," which appears to mean that they erred on at least one of the two problems.

The percent of ninth graders who showed weaknesses on such problems ranged from 47% to 94%; performance of college students in Guiler (1946a) was only slightly higher. Both the ninth graders and college students solved percent-whole items most accurately, although almost half of the students in both populations showed weaknesses on them.

More recent data from the National Assessment of Educational Progress (NAEP) indicate that the situation has not changed much in the past 75 years. The NAEP is a standardized test presented to nationally representative samples of US students every 2 or 4 years. Its mathematics subtest is given to roughly 150,000 fourth graders and 150,000 eighth graders on each testing occasion. The children who received the items on percentages described in this section were eighth graders (typically 13- or 14-year-olds) whose formal instruction in percentages had been completed.

The NAEP periodically releases a subset of items that have been used on previous tests, to provide information about the types of items that are presented. These released items provide a useful context for understanding the NAEP results. On the 11 released items that tested eighth graders' knowledge of percentages on NAEP exams between 1990 and 2017, mean percent correct was roughly 40% (U.S. Department of Education, 1990–2017). This level of accuracy was higher than the 20% correct that would have been expected by chance on these five-choice items, but far from mastery. Despite numerous reform efforts intended to improve understanding of rational numbers, including the Common Core State Standards

(National Governors Association Center for Best Practices, 2010), there were no obvious trends over time in accuracy on these released items.

To appreciate what this level of accuracy means, consider the following three items, the response alternatives available on each, and the percent correct on each:

"Of the following, which is the closest approximation of a 15% tip on a restaurant check of $24.99?" The response alternatives were $2.50, $3.00, $3.75, $4.50, and $5.00. Only 38% of the eighth graders chose the correct answer (NAEP, 1996).

"There were 90 employees in a company last year. This year, the number of employees increased by 10%. How many employees are in the company this year?" With response alternatives 9, 81, 91, 99, and 100, only 37% of students answered correctly (NAEP, 2005).

"Jared wants to buy a jacket that has a price of $49.99. He has a coupon for a discount of 30% of the price of the jacket. What is the price of the jacket after the 30% discount?" Given the choices of $14.99, $19.99, $34.99, $48.99, and $49.99, 47% of students chose the correct answer (NAEP, 2017).

Performance on these NAEP questions, together with Guiler's (1946a) findings, indicate that middle school children's understanding of percentages has been weak for at least 75 years. Moreover, many contemporary university students show similar lack of understanding. Jacobs Danan and Gelman (2017) presented students at a selective university with two-step problems involving either an increase of a given percentage from a base and then a decrease of the same percentage from the new value, or an increase of a given percentage and then a decrease of a different percentage. Some problems were presented with numerical values specified, others in algebraic form. Accuracy varied with problem features, but even the easiest types of problems elicited only 75% correct answers, whereas the hardest problems elicited 38% correct answers. Thus, U.S. students' weakness in understanding percentages is apparent even among students at selective universities.

4. When and why are percentages used?

4.1 Quantification process theory

Fractions, decimals, and percentages have all been in widespread use for at least three centuries (Cajori, 1993). The advantages of fractions are straightforward; they alone can express all rational numbers precisely. The advantages of decimals also are straightforward; they are a direct extension of the base-10 system used with whole numbers, allow straightforward mapping onto the metric system of measurement, and allow both children and adults to estimate magnitudes more accurately than they can with fractions (Hurst & Cordes, 2016, 2018). But why do we also use percentages? All percentages are equivalent to two-digit decimals; why have a separate notation for them?

To answer this question, Tian, Braithwaite, and Siegler (2020a) proposed *quantification process theory*. A specific instance of overlapping waves theory (Shrager & Siegler, 1998; Siegler, 1996), quantification process theory provides an explanation of how people choose among fractions, decimals, and percentages to represent proportional relations, for example in visual displays such as those in Fig. 2.

The basic assumption of quantification process theory is that when people choose among rational number notations for representing proportional relations, their choice is in large part determined by the type of quantification process—counting, measuring, or estimating—that they use to quantify the proportions. When they use counting, they prefer fractions. When they use estimation, they prefer percentages. When they use measurement, especially with metric measures, they prefer decimals.

Underlying these predictions about choices among quantification strategies are general principles of overlapping waves theory. This theory proposes that strategy choices are determined by the accuracy and speed

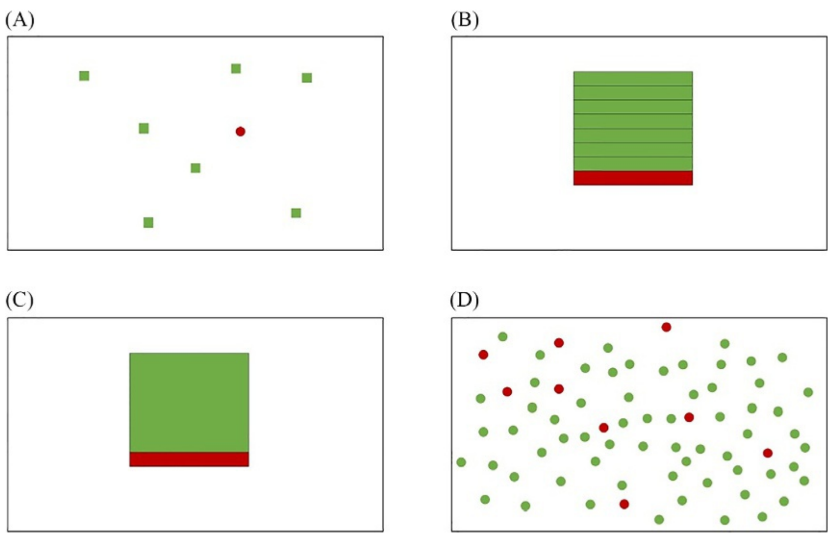


Fig. 2 Spatial displays representing (A) small-number discrete, (B) small-number discretized, (C) continuous, and (D) large-number discrete proportions.

yielded by strategies on specific problems, on problems with particular features, and on problems in the general domain. The cognitive cost of executing the strategies is also hypothesized to influence strategy choices. When the accuracy of different strategies varies, it usually is weighed most strongly in the choice process. In choosing among rational number notations, this leads to a preference for using fractions on the types of discrete and discretized displays shown in Fig. 2A and B, because counting yields exact values of numerators and denominators and thus fractions that are exactly correct. Note, however, that this prediction holds true only if the numbers of objects in the display, available time, and instructions allow use of counting. If there are too many objects to count in the available time, or if instructions indicate that counting should not be used, quantification process theory predicts that fractions will not be favored to represent discrete or discretized displays.

The role of cognitive demands within overlapping waves theory suggested that when people estimate proportions within displays, they should prefer percentages over decimals. Percentages have a fixed implicit denominator; each percentage is relative to 100. This fixed implicit denominator limits choices among percentages to 101 values. In contrast, decimals can be specified relative to any power of 10; thus, the same decimal can be described as 0.7, 0.73, 0.732, etc. The lack of a fixed denominator for decimals requires choosing a level of precision as well as a value within that level of precision. Choosing the level of precision adds to the cognitive demand of using decimals to describe proportions.

In cases where precision beyond the nearest percent is needed, for example in timing Olympic races or weighing precious metals, measurement instruments such as high precision timers and scales are typically used as the means of quantification. This leads to decimals typically being used to express the value. Neither estimation nor counting can be used in such situations, because they do not yield sufficiently precise results. Thus, when the goal is to measure a continuous dimension to a high degree of precision, decimals tend to be preferred.

4.2 Tests of quantification process theory

Quantification process theory was formulated in part as an alternative to a prior theory that sought to answer similar questions, *semantic alignment theory* (DeWolf, Bassok, & Holyoak, 2015). In the context of rational number notations, semantic alignment theory focuses on the choice between fractions and decimals to represent proportions in displays like those in Fig. 2A–C. It posits that the bipartite structure of fractions makes them inherently better suited to represent discrete displays, such as those in Fig. 2A and B, but that the unidimensional structure of decimals makes them inherently better suited to represent continuous displays, such as that in Fig. 2C.

Consistent with this perspective, most participants in studies testing semantic alignment theory have chosen decimals to represent continuous displays and fractions to represent discrete and discretized displays. This pattern of choices has been found not only among U.S. participants (DeWolf et al., 2015) but also in Korean and Russian samples (Lee, DeWolf, Bassok, & Holyoak, 2016; Tyumeneva et al., 2018).

Percentages have only been examined in one study based on semantic alignment theory (Gray, DeWolf, Bassok, & Holyoak, 2017). Gray et al. (2017) proposed that percentages have a unidimensional structure like that of decimals, and therefore hypothesized that when the choice was between percentages and fractions, percentages would be preferred to represent continuous displays, and fractions would be preferred to represent discrete displays. This prediction proved accurate.

Semantic alignment and quantification process theories make the same predictions for preferences between fractions and decimals or between fractions and percentages on the types of displays shown in Fig. 2A–C. The logic underlying the predictions differs, but the predictions are the same. However, the two theories lead to different predictions under at least three conditions.

(1) Quantification process theory predicts that percentages will be preferred over both decimals and fractions to represent continuous displays. The reason is that estimation will be used to quantify proportions on continuous displays and choosing a percentage to label the proportion is less taxing than choosing a decimal. Semantic alignment theory does not distinguish between decimals and percentages, because both are unidimensional notations, so it does not make this prediction.

(2) When large numbers of discrete objects are presented (Fig. 2D), quantification process theory predicts that percentages will be preferred when precision to the nearest percent is sufficient, because people will use estimation to approximate the proportions. In contrast, semantic alignment theory predicts that fractions will be used, because the bipartite nature of fraction notation matches the bipartite structure of the proportion being represented.

(3) When a discrete display is presented too briefly to allow accurate counting, quantification process theory predicts that participants will estimate and therefore choose percentages over fractions and decimals to represent the proportional relation. Semantic alignment theory makes the opposite prediction, because the objects are discrete regardless of the quantification process.

These predictions were tested in four experiments. Experiment 1 of Tian et al. (2020a) was designed to test the prediction from quantitative process theory that when estimation would be used to quantify the proportion illustrated in a display, participants would prefer percentages to both decimals and fractions. The three types of displays shown in Fig. 2A–C were presented to participants, who were students at a highly selective university. The displays were labeled small-number-discrete, small-number-discretized, and continuous. Small-number discrete and discretized displays involved between 7 and 13 units.

There were two experimental conditions. One condition was identical to that used in DeWolf et al. (2015); participants were asked whether to choose decimals or fractions to represent each display. In the other condition, participants were presented a new three-choice version of the task, in which the alternative notations included percentages as well as fractions and decimals, Following DeWolf et al. (2015), the task in both conditions involved choosing "which notation is the (better/best) representation of the depicted relation."

When percentages were not an option, choices paralleled those in the corresponding condition in DeWolf et al. (2015). Participants chose fractions as the desired rational number format on about 75% of trials for the discrete and discretized displays, and they chose decimals on about 75% of trials for the continuous displays. Choices in the three-choice condition also were similar for the small-number discrete and discretized displays. However, when presented the type of continuous displays shown in Fig. 2C, participants who had the option of choosing percentages did so overwhelmingly. To represent these continuous displays, only 15% of participants preferred decimals, vs 75% who preferred percentages.

At the insistence of an extraordinarily conscientious reviewer, Tian et al. (2020a) tested the possibility that the different findings with continuous displays might be due to one condition involving three response alternatives and the other two, as opposed to percentages being an option in the three-choice but not in the two-choice condition. To test this possibility, Tian et al. (2020a) presented additional two-choice conditions where one choice was between fractions and percentages and the other choice was between decimals and percentages. As expected, percentages were strongly preferred for representing continuous displays under both conditions: 82% of trials when the choice was between percentages and fractions and 75% of trials when the choice was between percentages and decimals. Thus, consistent with quantification process theory, percentages were preferred to decimals as well as fractions for representing proportions in continuous displays.

Experiment 2 of Tian et al. (2020a), tested the second hypothesis above, that when the number of discrete objects is large and the time to choose a notation limited, people prefer to quantify by estimating rather than by counting, and therefore would prefer percentages over fractions and decimals. We conducted this test by presenting both small-number discrete collections (6–13 objects) and large number discrete collections (70–91 objects) under the same two-choice and three-choice conditions as in Experiment 1. The experiment was conducted at a university in China, thus also allowing examination of whether notation preferences varied between societies with higher and lower proficiency with rational numbers (Bailey et al., 2015).

As predicted by quantification process theory, on the problems with large numbers of discrete objects, the Chinese students usually preferred percentages to express the proportion of objects that were a given color (about 60% of trials). Thus, it was the quantification process that determined the preferred notation, rather than whether the choice involved discrete or continuous proportions. When the number of objects was too large to allow rapid counting, students preferred estimation and therefore chose percentages to represent the proportion. The similarity of choices among students in the U.S. and China suggested that determinants of preferences among rational number notations are the same in societies with higher and lower levels of knowledge about rational numbers.

Experiment 3 of Tian et al. (2020a) tested two other predictions of quantification process theory. One was that the greater the number of objects to be quantified, the more often people would choose percentages, rather than fractions or decimals, to represent proportional relations within displays. The logic was that the more numerous the number of objects in the display, the more time that counting would take, and therefore the more likely people would quantify the proportional relation through estimation, which would lead to their choices of percentages increasing with increases in the number of objects.

The other prediction was that directly instructing participants to count would lead to them more often choosing fractions to represent proportions, regardless of the number of objects in the display, and that instructing them to estimate would lead to them more often using percentages to represent the same proportions, again regardless of the number of objects in the display. This was the most direct test of quantification process theory—changing the quantification process would change the preferred notation.

To test theses hypotheses, participants were presented 24 displays with 31–76 dots, some red and some green, and asked to generate a fraction, decimal, or percentage to represent the proportion of red dots in each display. When participants were asked to count, the display remained visible until the participant answered. Time to complete each trial was unlimited, to allow counting in that condition. In contrast, when participants were asked to estimate, the displays disappeared after 2s, to prevent counting.

The results supported both predictions from quantification process theory. The greater the number of dots in the display, the more often participants preferred percentages; the fewer dots in the display, the more often they preferred fractions.

Especially important for testing quantitative process theory, instructions to count or estimate greatly influenced choice of notation. When instructed to count, participants chose fractions to represent the proportion on 77% of trials; when instructed to estimate, participants chose percentages on 74% of trials. These findings clearly demonstrated the linkage between the quantification process and the choice of rational number notation.

In Experiment 4, Tian et al. (2020a) tested the hypothesis that the complexity of processing decimals with varying numbers of decimal digits led people to prefer percentages to decimals when precision beyond the nearest percent was unnecessary. They presented university students displays with red dots and green dots, like those in Fig. 2, and a pair of rational numbers: two fractions, two decimals, or two percentages. One number in each pair was very close to the exact proportion within the display; the other number differed by roughly 0.15 from the correct proportion. The task was to choose the number that better represented the proportion of dots of the desired color. On choices between decimals, the two alternatives always had different numbers of decimal digits, with the number varying between one and three decimal digits (thus, the choice might be between 0.427 and 0.57).

Patterns of accuracy on this task matched patterns of preferences in previous experiments. When large-number discrete displays were presented, choices between percentages were consistently more accurate than choices between decimals or fractions. These findings were consistent with the view that the greater cognitive load imposed when decimal digits vary contributes to preferences for percentages over decimals.

In sum, the findings of Tian et al. (2020a) were consistent with the predictions of quantification process theory. When estimation is used to quantify proportions, people generally prefer percentages over both decimals and fractions. When counting is used to quantify proportions, fractions are generally preferred to percentages and decimals.

Decimals were never the notation of choice in any of the four experiments in Tian et al. (2020a). This raised the issue of when decimals are preferred. One likely answer is that decimals are preferred in situations when precision of measurement beyond the nearest percent is important, as when timing races. Decimals also are likely to be preferred in contexts using metric measurements, due to the ease of using decimal notation with metric measures. Rapp, Bassok, DeWolf, and Holyoak's (2015) examination of textbook problems yielded results consistent with this hypothesis. On textbook problems using metric measures (e.g., centimeters, kilograms, liters), decimals were most often used; on textbook problems using imperial units (e.g., inches, pounds, quarts), fractions were most often used. Whether people prefer decimals in situations involving metric measures and fractions in situations using imperial units remains to be tested.

5. Textbook coverage of percentages

Data from the NAEP released items and other sources left little doubt that many students have only weak understanding of percentages. To develop hypotheses about the sources of this weak understanding, Tian (2018) examined coverage of percentages in two popular textbook series: Harcourt's *Go Math!* (Dixon, Adams, Larson, & Leiva, 2012) and Pearson's *enVisionmath* (Charles et al., 2012). Analyses of textbook input have proved useful for identifying the origins of a variety of difficulties in children's math learning, including difficulties with mathematical equality (McNeil, Fyfe, & Dunwiddie, 2015), order of operations (Landy & Goldstone, 2007), fraction arithmetic (Braithwaite, Pyke, & Siegler, 2017), and decimal arithmetic (Tian, Braithwaite, & Siegler, 2020b). The hope was that analyzing distributions of percentage arithmetic problems in textbooks would prove similarly illuminating.

Examination of these textbook series revealed several striking ways in which the coverage of percentages differed from that of fractions and decimals. One was that percentages were not the focus of even a single chapter in either textbook series before 6th grade. Another striking feature was that percentages received far less total coverage than fractions or decimals; whereas fractions were covered in 12 chapters in one textbook series and 14 in the other, and decimals were covered in 6 chapters in one series and 10 in the other, percentages were the focus of only 2 chapters in each series, one in 6th grade and one in 7th grade.

Altogether, in the 6th and 7th grade volumes of the two series, there were 802 problems involving percentages: 328 in *Go Math!* and 474 in *enVision Math*. Unlike in the previous analyses of textbook problems by Braithwaite et al. (2017) with fractions and Tian et al. (2020b) with decimals, this analysis of textbook problems with percentages included word problems as well as purely numerical problems, due to the high frequency of percentage word problems relative to purely numerical problems (e.g., "What is 70% of 50?"). The main types of percentage problems involved translation and arithmetic (Table 1).

5.1 Translation problems

One common type of problem assessed understanding of percentages as individual numbers. Some such items required translation between another type of number and a percentage (e.g., "Write 4/5 as a percent"). Others required translation between a percentage and a pictorial model (e.g., "Shade 14% of this 10×10 grid"). Problems requesting translations between percentages and other numerical notations were very frequent in 6th grade but rare in 7th grade. In *Go Math*, 69% of problems in the 6th grade textbook and 6% of the problems in the 7th grade textbook volume involved translation; in *enVision Math*, 49% of problems in the 6th grade volume and 6% in the 7th grade textbook did.

5.2 Arithmetic problems

The other main type of item involved arithmetic with percentages. In contrast to the translation problems, percentage arithmetic problems were common in both grades in both textbooks. In 6th grade, 47% of *Go Math!* and 56% of *enVision Math* problems involved percentage arithmetic; in 7th grade, 83% of *Go Math!* and 77% of *enVision Math* problems did. (The percentages of translation and arithmetic problems sometimes sum to more than 100% because some arithmetic problems also required translation from decimals or fractions to percentages. Other times, the percentage of problems sum to less than 100% because a few problems involved neither translation nor arithmetic.)

Table 1 Percentage of each type of problem in 6th and 7th grade volumes of Go Math! and enVision Math.

Content	**Description**	**Example**	**Grade 6 (*N* = 402)**	**Grade 7 (*N* = 400)**
Translate number	Translate among percentages, fractions, and decimals	Write 0.59 as a percent	41	6
Translate pictorial model	Translate between percentages and quantity represented by pictorial models	Shade the model to represent 14% [of a 10 × 10 blank grid]	15	1
100% as the whole	100% represents the whole	Many chemical elements can be found in Earth's atmosphere. What percent of gases in Earth's atmosphere does the whole circle graph represent? [pie chart colored blue for nitrogen, 78%; orange for oxygen, 21%, purple for other gases, 1%]	3	1
Percent–Part	Given the percent and the part, find the whole, or given the percent change and the new amount, find the original	A section of rope is 5 in. long, which represents 20% of the length of the entire rope. How long is the rope?	14	6
Whole–Part	Given the whole and the part, find the percent, or given the original and new amount, find the percent change	1. What percent of 30 is 24? 2. The price of a pair of shoes increases from $52 to $64. What is the percent increase to the nearest percent?	11	25

Continued

Table 1 Percentage of each type of problem in 6th and 7th grade volumes of Go Math! and enVision Math.—cont'd

Content	Description	Example	Grade 6 ($N=402$)	Grade 7 ($N=400$)
Percent-Whole	Given the whole and the percent, find the portion, or given the original amount and the percent change, find the new amount	1. Sharon wants to buy a shirt that costs \$20. The sales tax is 5%. How much is the sales tax? 2. A TV has an original price of \$499. Find the new price after a 30% decrease	24	49
Other	Problems that cannot be categorized		4	16

Note: Percentages in some columns sum to more than 100%, because some problems fit into more than one category.

Almost all percentage arithmetic problems were percent-whole-part problems, in which the problem specified two of the variables and students needed to compute the third. Some of these problems involved changes over time (e.g., "A TV has an original price of \$499. Find the new price after a 30% discount"), whereas other problems did not (e.g., "What is 31% of 94"). Both were grouped together in the present analyses according to the two variables that were specified in the problem.

The "whole $\times$ percentage $=$ part" problems, in which the problem statement indicated the whole and the percentage and the participant needed to provide the part, were most frequent; they appeared often in both textbook series in both 6th and 7th grade. In contrast, "part $\div$ percentage $=$ whole" problems were much less common. As in previous analyses of problem distributions for fractions (Braithwaite et al., 2017) and decimals (Tian et al., 2020b), relative frequency of problems was similar across different textbook series.

5.3 Differences between textbook coverage of arithmetic with percentages and with other types of rational numbers

Textbook coverage of percentage arithmetic was distinctive in at least four ways:

(1) Almost all percentage arithmetic problems involved multiplication (more than 90% in both textbook series). In arithmetic with fractions and decimals, no one operation predominated.
(2) Almost all percentage multiplication problems (more than 90% in both textbooks) involved a percentage and a whole number. This was unlike other textbook problems, which usually involved two numbers of the same type (Braithwaite et al., 2017; Tian et al., 2020b).
(3) Almost all (99%) problems involving multiplication of a percentage and a whole number were phrased as "N% *of* the whole number." In contrast, only 21% of fraction multiplication problems used the "of" phrasing (e.g., ¾ of 40), and 0% of the decimal multiplication problems did. All fraction problems that used the "of" phrasing were problems with fractions below one.
(4) Word problems constituted a large majority of problems involving percentage multiplication in both textbooks in both grades, whereas most problems involving multiplication with fractions and decimals just specified the operands and the arithmetic operation.

The greater frequency of the "of" phrasing, relative to the multiplication sign, on problems involving multiplication of a percentage and a whole number suggested that on such problems, accuracy would be higher with the "of" phrasing. Another reason for the same prediction was that the "of" phrasing seems more conceptually transparent (e.g., 75% of 40 seems more conceptually transparent than $75\% \times 40$). On the other hand, the multiplication sign might have been expected to elicit more accurate performance than the "of" phrasing, because the "of" phrasing requires translation of the "of" into the operation of multiplication.

In contrast, frequency of problems from the same textbooks suggested the opposite prediction for decimal multiplication. The multiplication sign was always used with decimal multiplication in both textbooks, leading to the prediction that problems with the multiplication sign would be solved more often than corresponding problems with the "of" phrasing.

The pattern with fraction multiplication was more difficult to anticipate. The greater frequency of the multiplication sign than the "of" phrasing suggested that performance with the multiplication sign would be superior. On the other hand, conceptual transparency suggested that the "of" interpretation was a more intuitive way to think about fraction multiplication than was use of the multiplication sign, at least with fractions below one (e.g., "3/4 of 40" vs "$3/4 \times 40$"). Whether that was true with fractions above one (e.g., "7/4 of 40") was less clear. The fact that 27% of multiplication problems in the two textbooks with fractions below one involved the "of" phrasing, vs 0% with fractions above one, also suggested that phrasing might interact with whether the fraction was above or below one.

5.4 Textbook problem distributions and children's performance

The hypothesis that multiplication problems involving a percentage and a whole number, and perhaps multiplication problems involving a fraction and a whole number, would benefit from use of the "of" phrasing, was tested by presenting direction of effect problems of the form "True or False: $N \times M > N$, with $N > M$." Middle school students have been found to usually answer incorrectly direction of effects multiplication problems involving numbers between 0 and 1 with two fractions (Fischbein, Deri, Nello, & Marino, 1985; Siegler & Lortie-Forgues, 2015) and two decimals (Lortie-Forgues & Siegler, 2017). Inaccuracy on such problems appears to reflect overgeneralization from whole number multiplication, where products equal or exceed both multiplicands.

Tian (2018) tested the impact of the "of" phrasing on direction of effects judgments on problems involving multiplication with a whole number and either a percentage, a fraction, or a decimal. More than 200 sixth and eighth graders were randomly assigned to solve problems involving fractions, decimals, and percentages. Students were first presented direction of effects multiplication problems with the multiplication sign (e.g., "True or false: $3/4 \times 40 > 40$") and then problems with the "of" phrasing (e.g., "True or false: $3/4$ of $40 > 40$"). Half of the problems in each set involved rational numbers greater than one, and half involved rational numbers less than one. In each problem, the first multiplicand was always a rational number (percentage, fraction, or decimal), and the second multiplicand was always a whole number. On corresponding problems, the rational number was equivalent or almost so; the same whole number was used with all three corresponding rational numbers. Participants were told not to calculate but rather just to judge whether each statement was true or false.

The data were consistent with the predictions of the above analysis. On direction of effects problems involving multiplication of a percentage and a whole number, the "of" phrasing elicited more accurate performance, regardless of whether the percentage was above or below 100% (Fig. 3). This was consistent with both textbook frequency and the view that the "of" phrasing was more transparent with percentages. On problems with decimals, where the "of" phrasing never appeared in textbooks,

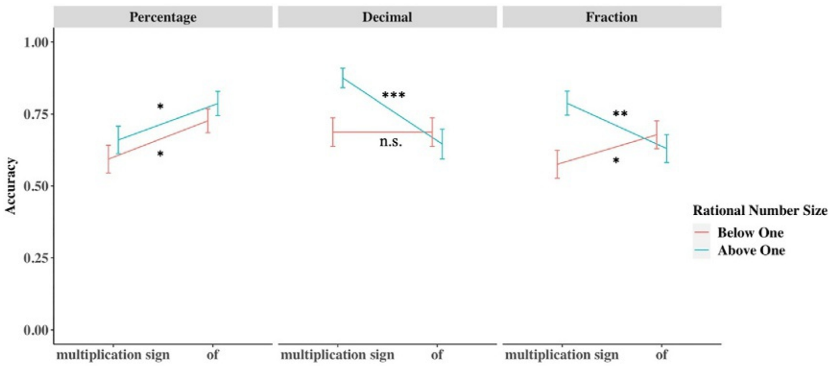


Fig. 3 Sixth and eighth graders' accuracy on direction of effects judgments of percentage, decimal, or fraction multiplication presented with the multiplication sign or the "of" phrasing with numbers above and below one. "n.s." indicates non-significant, "*" indicates $P<0.05$, "**" indicates $P<0.01$, and "***" indicates $P<0.001$.

that phrasing elicited less accurate performance than the multiplication sign. On problems with fractions, there was no overall effect of phrasing on accuracy of the direction of effect judgments, but this reflected an interaction between the phrasing and whether the fraction was below or above one. On problems where the fraction was above one, where the "of" phrasing never appeared in textbooks, judgment accuracy was less accurate with the "of" phrasing than with the multiplication sign. In contrast, on problems where the fraction was less than one, children's judgments were more accurate with the "of" phrasing.

These findings suggested that the effects of the "of" phrasing, relative to the multiplication sign, depended on both familiarity with the two ways of presenting the problems and on their conceptual transparency. The "of" phrasing was unambiguously helpful on direction of effects problems involving percentages, where both textbook frequency and conceptual transparency predicted more accurate judgments with the "of" phrasing. On problems where the "of" phrasing never appeared in textbooks, notably problems with decimals and with fractions above one, judgments were more accurate with the standard multiplication sign. However, when the "of" phrasing was used sometimes but not on a majority of trials (fractions below one), the "of" phrasing promoted more accurate performance than did the multiplication sign.

Thus, the "of" phrasing appears to increase the accuracy of direction of effects judgments on multiplication problems as long as that phrasing is at least somewhat familiar. If the phrasing is totally unfamiliar, however, the "of" phrasing seems to reduce the accuracy of judgments.

6. Estimating answers to percentage multiplication problems

Direction of effects problems are a useful index of conceptual understanding, but they have at least three limitations. One is that solutions to direction of effects problems can be learned as a simple rule: Multiplying numbers greater than one yields answers greater than either multiplicand; multiplying numbers between zero and one yields answers less than either multiplicand. Students might learn this rule either via direct classroom instruction or by generalizing from their experience with multiplication, without understanding why it holds.

A second, related, limitation of direction of effects tasks is that problems with rational number operands greater than one can be answered correctly by generalizing from whole number multiplication. Students might have no understanding of what percentages meant, but if they assumed that they worked like whole numbers, they would correctly answer direction of effects tasks involving multiplication of percentages greater than 100%.

A third limitation is that direction of effects tasks only yield information about whether judgments are accurate, rather than information about the degree of understanding. If two students answer such problems correctly, one might have an accurate sense of the products yielded by the multiplication whereas the other might only have a general sense that the answer is greater or less than the larger operand.

To provide a more nuanced assessment of conceptual understanding of percentage multiplication, and to test the generality of the beneficial effects of the "of" phrasing with percentages and fractions below one, Tian (2018) examined estimation of answers to percentage and fraction multiplication problems. A person with a good understanding of percentage multiplication might estimate that 25% of 33 would be about 8 or 9, whereas one who did not might estimate that the answer would be about 800 (because 25×33 is about 800.)

One purpose of this experiment was to examine the accuracy of such estimates. Another purpose was to test the generality of findings from the direction of effects task regarding the beneficial effects of the "of" phrasing on percentage and fraction multiplication. Would this phrasing improve the accuracy of quantitative estimates of multiplication involving percentages as it did the accuracy of the qualitative judgments previously? A third purpose was to examine the role of benchmarks in rational number calculation. In this context, benchmarks are numbers that facilitate approximation of answers to arithmetic problems due to inherent characteristics of the numbers, familiarity of the numbers, or both. Fractions less than one with a denominator of 2, 3, 4, or 10, and corresponding percentages, were labeled "benchmarks," because they seemed likely to facilitate accurate estimation of answers to multiplication problems.

To examine these issues, Tian (2018) presented pre-service teachers with problems involving multiplication of a whole number by a percentage or corresponding fraction. In one condition, participants were presented problems with the "of" phrasing; in the other condition, participants were presented problems with the multiplication sign. In both conditions, participants were asked to state an answer that was as close as possible to the correct answer within 8 s, the time limit being used to prevent exact calculation.

There were four types of problems each with eight items: fraction benchmark (e.g., $1/4 \times 82$), fraction non-benchmark (e.g., $6/23 \times 79$), percentage benchmark (e.g., $25\% \times 82$), and percentage non-benchmark (e.g., $26\% \times 79$). All fractions were between 0 and 1, and all percentages were between 0% and 100%, ranges in which the "of" phrasing had increased the accuracy of judgments in the direction of effects study. Each non-benchmark number was close, but not equal, to a paired benchmark (e.g., 1/3 vs 6/19, 90% vs 89%). The anticipated best answer was the same for the benchmark and corresponding non-benchmark problems (e.g., the answer would be "20" for both $25\% \times 82$ and $26\% \times 79$).

Accuracy on these problems involving multiplication of a whole number by a fraction or percentage was analyzed relative to both the correct answer and the answer yielded by the *optimal estimation strategy*. This optimal estimation strategy involved rounding the whole number to the nearest 10, rounding non-benchmark fractions and percentages to the nearest benchmark, and translating percentages that were not multiples of 10 to the nearest benchmark fraction. For example, on $76\% \times 38$ the optimal strategy involved rounding 76% to 75%, translating 75% to 3/4, rounding 38 to 40, and then multiplying "$3/4 \times 40$" to get 30. The answer yielded by the optimal estimation strategy was always within one of the correct product after rounding. For example, $76\% \times 38 = 28.88$, which when rounded to the nearest whole number, 29, was within one of 30. Answers yielded by the optimal estimation strategy and answers correct to the nearest whole number were both classified as accurate.

To obtain a general sense of the quality of the pre-service teachers' answers, we first examined the percentage of answers that met this accuracy criterion. Overall, 49% of answers met it. The "of" phrasing led to a greater percent correct than did the multiplication sign (51% vs 47%). Problems with benchmark fractions or percentages more often elicited accurate answers than did problems with non-benchmark fractions or percentages (59% vs 39%). Fractions and percentages led to similar frequency of accurate answers (48% and 49%). A linear mixed effects regression analysis yielded support for each of these conclusions (see Tian, 2018, for details).

Tian (2018) also examined the frequency with which answers met the direction-of-effects criterion for accuracy. Because all fractions were below one and all percentages were below 100%, answers that met this criterion were ones where the answer was smaller than the whole number operand. A mixed-effects logistic regression of the relation between phrasing, benchmark, rational number notation (fraction or percentage), and correctness of direction of effects was conducted. We also entered a random effect of problem. The direction of effects criterion was more often met on problems with the "of" phrasing (95%; 97% for percentages and 93% for fractions) than on ones with the multiplication sign (83%; 84% for percentages and 82% for fractions), $\beta = -1.85$, $P < 0.001$. This finding was consistent with the view that the "of" phrasing promotes conceptual understanding of percentage and fraction multiplication.

7. Instructional implications

What type of instruction would improve learning about percentages, as well as other types of rational numbers? Perhaps the best evidence regarding this question comes from studies of a curriculum developed by Moss (1997), Moss & Case, 1999), which started with percentages, used learning of them to promote learning about decimals and fractions, and then focused on connections among the three rational number notations. The theory underlying the curriculum resembled the integrated theory in emphasizing the importance of numerical magnitudes, number lines, and commonalities among different types of rational numbers. The instruction also built on children's prior knowledge of spatial proportions, doubling, halving, and additive composition (the concept that numbers are combinations of other numbers).

To test the effectiveness of the curriculum, Moss and Case (1999) presented 4th graders, who were described as having no prior instruction in rational numbers, either the experimental curriculum or a control curriculum that was widely used in Canada where the study was conducted. Both curricula involved extensive instruction—25 sessions in the experimental curriculum, 20 in the control curriculum.

The initial sessions of the experimental curriculum built on children's knowledge of whole numbers akin to the percentages 0–100, as well as their knowledge of percentages from everyday life. On some problems, children were presented cylindrical beakers and asked to draw beakers with levels of liquid and percentages that indicated their fullness (e.g., "Draw this cylinder when it's 25% full"). On other problems, the children were asked to estimate the percentage fullness of a beaker. This approach was consistent with findings that by fourth grade, children represent the numbers 0–100 linearly (Siegler & Opfer, 2003), a representation that could be applied straightforwardly to percent fullness judgments with the cylindrical beakers.

Children were also encouraged to use strategies consistent with the representations that the curriculum was trying to build. On such strategy was halving to assign numbers to benchmark water levels, such as indicating levels of liquid that were 100%, 50%, and 25% full. Another was the strategy of composition, which involved adding percentages in the same way as whole numbers, both numerically and spatially (e.g., "How full would a beaker be if we added the water in a 75% full beaker to that in a 25% full beaker?").

Later, children were encouraged to extend the knowledge they had gained about percentages to two-digit decimals. They were told that the decimals indicated the percent distance between two whole numbers, for example, 5.25 representing a location 25% of the distance between 5 and 6. The idea was eventually generalized to decimals with other numbers of decimal digits. This part of the curriculum led some children to invent an interesting notation, the "double decimal," in which numbers such as 5.25.25 were viewed as representing 25% of the distance from 5.25 to 5.26. Although this construction deviates from standard decimal notation, it reflects the success of Moss' curriculum in motivating children to reason about rational numbers rather than just memorizing what they are told. Fractions were used at a variety of times in the curriculum, at first informally with common terms such as "half full," later in having students translate fractions into equivalent spatial displays, and yet later in more formal ways.

The final part of the curriculum presented exercises that connected percentages, decimals, and fractions, an approach that not only promoted an integrated number sense but also conveyed the concept of numerical equivalence. For example, students were asked to translate fractions into equivalent decimals and percentages. Rather than introducing fractions, decimals, and percentages in three different grades, all three notations were taught contiguously within a relatively short time in a single unified framework. Such contiguous instruction seems useful for avoiding the siloed knowledge of rational numbers that often emerges when fractions, decimals, and percentages are focused on in different grades.

Children's knowledge improved from pretest to posttest in both conditions. However, the progress was considerably greater among children who received the experimental curriculum. This pattern was present for knowledge of percentages, decimals, and fractions considered separately, as well as for knowledge of rational number magnitudes ("Draw a picture to show which is greater, 2/3 or 3/4"), translation across notations ("What is 1/8 as a decimal"), density ("Is there a number between 0.3 and 0.4; can you name one"), computation ("What is 65% of 160"), and solutions to word problems ("These CD's are marked down from \$8.00 to \$7.20. What is the discount as a percentage of the original price?"). Differences between the experimental and control groups in the percent of children correctly solving the problems on the posttest tended to be very large. For the five problems cited above, percent correct on the posttest among children who were presented the experimental and control curricula were 81% vs 38%, 75% vs 0%, 100% vs 15%, 69% vs 0%, and 56% vs 8%. A subsequent study (Kalchman, Moss, & Case, 2001) replicated these findings regarding the benefits of Moss's experimental curriculum with a different group of 4th graders and a different control curriculum, and among 6th graders as well.

Moss's curriculum differed from the control curriculum in many ways, including order of instruction about the three rational number notations, emphasis on halving and composition as problem solving strategies, frequency of use of number lines to illustrate magnitude relations, emphasis on relations among the three notations, and relative emphasis on conceptual understanding vs procedural mastery. These and other differences precluded isolation of the contribution of each factor in the experimental curriculum to its overall effectiveness.

We suspect, however, that presenting percentages first, and using knowledge of percentages to facilitate understanding of the other two notations, played a particularly important role. This speculation is based on four lines of reasoning. One is that percentages are easy to map onto whole numbers in the 0–100 range. Just as 80 > 40, so is 80% > 40%. Moreover, ratio relations are maintained; 80% is twice as large as 40%, just as 80 is twice as large as 40. These similarities seem like a large advantage in helping children learn about rational numbers.

A second advantage of starting rational number instruction with percentages, rather than fractions or decimals, is that children seem to encounter percentages more often in everyday life. Situations in which children often encounter percentages, such as interpreting test scores, buying goods on sale, and noting charges on electronic devices, are important to children, which may further enhance their informal learning about percentages.

A third line of reasoning regarding the potential advantages of teaching percentages first involves cognitive load. The implicit denominator of percentages is fixed; it is always 100. In contrast, the implicit denominator of decimals can be any power of 10, and the denominator of fractions can be any number. Variable denominators seem likely to increase the processing demands of representing magnitudes. Consistent with this analysis, adults were more accurate in choosing percentages than decimals to label proportions when the number of decimal digits varied, but accuracy was equal for percentages and decimals when the number of decimal digits was always two (Gray et al., 2017; Tian et al., 2020b).

A fourth line of reasoning, perhaps the most important, involves helpful and harmful analogies. As in other domains, people draw both helpful and harmful analogies in learning about rational numbers. One common helpful analogy is that just as whole numbers have magnitudes that can be placed and ordered on number lines, so do fractions, decimals, and percentages. One common harmful analogy is that just as there are no whole numbers between 3 and 4, there are no decimals between 0.3 and 0.4 or fractions between 2/5 and 3/5. The relative strengths of helpful and harmful analogies considerably influence learning about rational numbers (Braithwaite et al., 2017).

The most frequent analogies that children draw in learning about rational numbers are from whole numbers. Relating percentages to whole numbers in the way that Moss's curriculum did seems particularly likely to promote helpful analogies and avoid harmful ones. Encouraging children to think of the fullness of the beakers in terms of percentages seems likely to have promoted a helpful analogy between percentages and whole numbers between 0 and 100. Halving and composition work the same ways with percentages as with whole numbers; encouraging those strategies again promotes a useful analogy.

The relation of whole numbers to fractions and decimals is more complex, Harmful analogies with whole numbers are harder to avoid, and the harm often persist for years. As late as 8th grade, many children estimate fractions with larger whole number components as having greater magnitude than equivalent fractions with smaller components (e.g., they see 6/9 as larger than 2/3) (Braithwaite & Siegler, 2017). Similarly, both children and adults more rapidly compare decimal magnitudes when the number of decimal digits is greater for the larger number than when the opposite is the case (Ren & Gunderson, 2019; Varma & Karl, 2013). Percentages do not seem to encourage similarly misleading analogies.

In the U.S. and many other countries, fractions are taught before decimals, and decimals are taught before percentages. The reason for this ordering is unclear, though. Learning outcomes certainly are not the reason; understanding fractions, in particular, is limited among students throughout the world (Lortie-Forgues, Tian, & Siegler, 2015).

One rationale for teaching fractions first might be that they are a simple extension of whole number division: $7/4 = 7 \div 4$. The extension is straightforward, however, only with fractions greater than one; many students have no idea why $4/7 = 4 \div 7$, or even what $4 \div 7$ means. Moreover, understanding of whole number division is often limited to execution of procedures among both teachers (Ma, 1999) and children (Dubé & Robinson, 2018). Thus, understanding of whole number division seems a slim reed on which to justify instructing students in fractions before other rational number notations.

Inverting this order and teaching percentages first, then decimals, and then fractions might yield better learning of all three notations. The advantages of starting with percentages have already been noted. Decimals, like percentages, are straightforward extensions of the place value system used with whole numbers, and prior learning of percentages can be extended to them, as illustrated in the findings of Moss and colleagues. Moreover, Moss and colleagues' findings indicated that the sequence of percentages followed by decimals yielded superior learning of fractions as well as decimals and percentages. The case for teaching percentages first is far from definitive; however, it seems well worth additional tests.

8. Conclusions

The findings and observations in this chapter indicate that percentages are widely used in everyday life, often preferred over fractions and decimals, poorly understood by many US students, and minimally emphasized in textbooks. The findings also demonstrate that almost all percentage arithmetic problems in current textbooks involve multiplication and that the multiplication almost always involves a percentage and a whole number. The difficulty that many children encounter with percentages seems especially striking given the narrow range of problems they are asked to solve.

Although difficulties learning percentages have persisted among U.S. students for at least 75 years, the difficulties seem far from inevitable. Understanding can be improved through small changes in problem presentation, such as using the "of" phrasing for multiplication. Larger improvements can be produced by large-scale interventions, such as that of Moss and colleagues. The integrative theory of numerical development, as well as Moss' findings and those of small-scale interventions (e.g., Braithwaite & Siegler, 2021; Fazio et al., 2016), suggest that emphasizing numerical magnitudes, number line representations, and interconnections among numbers can improve understanding of percentages and other types of rational numbers as well. We can hope that successful tests of these ideas, and implementation of the successful approaches in classrooms, will allow a quite different depiction of development of understanding of percentages 75 years from now.

References

Bailey, D. H., Siegler, R. S., & Geary, D. C. (2014). Early predictors of middle school fraction knowledge. *Developmental Science*, *17*(5), 775–785. https://doi.org/10.1111/desc.12155.

Bailey, D. H., Zhou, X., Zhang, Y., Cui, J., Fuchs, L. S., Jordan, N. C., et al. (2015). Development of fraction concepts and procedures in U.S. and Chinese children. *Journal of Experimental Child Psychology*, *129*, 68–83. https://doi.org/10.1016/j.jecp.2014.08.006.

Booth, J. L., & Siegler, R. S. (2008). Numerical magnitude presentations influence arithmetic learning. *Child Development*, *79*(4), 1016–1031. https://doi.org/10.1111/j.1467-8624.2008.01173.x.

Braithwaite, D. W., Pyke, A. A., & Siegler, R. S. (2017). A computational model of fraction arithmetic. *Psychological Review*, *124*(5), 603–625. https://doi.org/10.1037/rev0000072.

Braithwaite, D. W., & Siegler, R. S. (2017). Developmental changes in the whole number bias. *Developmental Science*, *124*, 603–625. https://doi.org/10.1111/desc.12541.

Braithwaite, D. W., & Siegler, R. S. (2018). Children learn spurious associations in their math textbooks: Examples from fraction arithmetic. *Journal of Experimental Psychology: Learning Memory and Cognition*, *44*(11), 1765–1777. https://doi.org/10.1037/xlm0000546.

Braithwaite, D. W., & Siegler, R. S. (2021). Putting fractions together. *Journal of Educational Psychology*, *113*, 556–571. https://doi.org/10.1037/edu0000477.

Braithwaite, D. W., & Siegler, R. S. (2021). Putting fractions together. *Journal of Educational Psychology*, *113*(3), 556–571. https://doi.org/10.1037/edu0000477.

Braithwaite, D. W., Tian, J., & Siegler, R. S. (2018). Do children understand fraction addition? *Developmental Science*, *21*(4). https://doi.org/10.1111/desc.12601, e12601.

Cajori, F. (1993). *A history of mathematical notations*. Dover Publications Inc. https://doi.org/10.1038/125078a0.

Carey, S. (2009). Where our number concepts come from. *Journal of Philosophy*, *106*(4), 220–254.

Castronovo, J., & Göbel, S. M. (2012). Impact of high mathematics education on the number sense. *PLoS One*, 7(4), 24–25. https://doi.org/10.1371/journal.pone.0033832.

Charles, R., Caldwell, J., Cavanagh, M., Chancellor, D., Copley, J., Crown, W., et al. (2012). *enVisionmath (common core ed.)*. Pearson Education, Inc.

DeWolf, M., Bassok, M., & Holyoak, K. J. (2015). Conceptual structure and the procedural affordances of rational numbers: Relational reasoning with fractions and decimals. *Journal of Experimental Psychology. General*, *144*(1), 127–150. https://doi.org/10.1037/xge0000034.

Dixon, J. K., Adams, T. L., Larson, M., & Leiva, M. (2012). *GO MATH! (common core ed.)*. Houghton Mifflin Harcourt Publishing Company.

Dubé, A. K., & Robinson, K. M. (2018). Children's understanding of multiplication and division: Insights from a pooled analysis of seven studies conducted across 7 years. *British Journal of Developmental Psychology*, *36*(2), 206–219. https://doi.org/10.1111/bjdp.12217.

Fazio, L. K., Kennedy, C. A., & Siegler, R. S. (2016). Improving children's knowledge of fraction magnitudes. *PLoS One*, *11*(10), 1–14. https://doi.org/10.1371/journal.pone.0165243.

Fias, W., Sahan, M. I., Ansari, D., & Lyons, I. M. (2021). From counting to retrieving: Neural networks underlying alphabet arithmetic learning. *Journal of Cognitive Neuroscience*, *34*(1), 16–33. https://doi.org/10.1162/jocn_a_01789.

Fischbein, E., Deri, M., Nello, M. S., & Marino, M. S. (1985). The role of implicit models in solving verbal problems in multiplication and division. *Journal for Research in Mathematics Education*, *16*(1), 3–17. https://doi.org/10.2307/748969.

Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities*, *37*(1), 4–15. https://doi.org/10.1177/00222194040370010201.

Geary, D. C. (2006). Development of mathematical understanding. In *Cognition, perception and language: Vol. 2. Handbook of child psychology* (pp. 777–810). https://doi.org/10.1002/9780470147658.chpsy0218.

Gelman, R. (1991). Epigenetic foundations of knowledge structures: Initial and transcendent constructions. In S. Carey, & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition* (pp. 293–322). Erlbaum.

Gelman, R., & Williams, E. M. (1998). Enabling constraints for cognitive development and learning: Domain-specificity and epigenesis. In *Cognition, perception, and language: Vol. 2. Handbook of child psychology* (pp. 575–630).

Gray, M. E., DeWolf, M., Bassok, M., & Holyoak, K. J. (2017). Dissociation between magnitude comparison and relation identification across different formats for rational numbers. *Thinking & Reasoning*, *6783*(October), 1–19. https://doi.org/10.1080/13546783.2017.1367327.

Guiler, W. S. (1946a). Difficulties encountered in percentage by college freshmen. *Journal of Educational Research*, *40*(2), 81–95. http://www.jstor.org/stable/27528757.

Guiler, W. S. (1946b). Difficulties in percentage encountered by ninth-grade pupils. *The Elementary School Journal*, *46*(10), 563–573. https://doi.org/10.1086/458886.

Gunderson, E. A., Ramirez, G., Beilock, S. L., & Levine, S. C. (2012). The relation between spatial skill and early number knowledge: The role of the linear number line. *Developmental Psychology*, *48*(5), 1229–1241. https://doi.org/10.1037/a0027433.

Handel, M. (2016). What do people do at work? A profile of U.S. jobs from the survey of workplace skills, technology, and management practices (STAMP). *Journal for Labour Market Research*, *49*(2), 177–197. https://doi.org/10.1007/s12651-016-0213-1.

Hurst, M. A., & Cordes, S. (2016). Rational-number comparison across notation: Fractions, decimals, and whole numbers. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(2), 281–293. https://doi.org/10.1037/xhp0000140.

Hurst, M. A., & Cordes, S. (2018). A systematic investigation of the link between rational number processing and algebra ability. *British Journal of Psychology*, *109*(1), 99–117. https://doi.org/10.1111/bjop.12244.

Jacobs Danan, J. A., & Gelman, R. (2017). The problem with percentages. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1740), 20160519. https://doi.org/10.1098/rstb.2016.0519.

Kalchman, M., Moss, J., & Case, R. (2001). Psychological models for the development of mathematical understanding: Rational numbers and functions. In S. Carver, & D. Klahr (Eds.), *Cognition and instruction: Twenty-five years of progress* (pp. 1–38). Erlbaum.

Landy, D., & Goldstone, R. L. (2007). How abstract is symbolic thought? *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *33*(4), 720–733. https://doi.org/10.1037/0278-7393.33.4.720.

Lee, H. S., DeWolf, M., Bassok, M., & Holyoak, K. J. (2016). Conceptual and procedural distinctions between fractions and decimals: A cross-national comparison. *Cognition*, *147*, 57–69. https://doi.org/10.1016/j.cognition.2015.11.005.

Linsen, S., Verschaffel, L., Reynvoet, B., & De Smedt, B. (2015). The association between numerical magnitude processing and mental versus algorithmic multi-digit subtraction in children. *Learning and Instruction*, *35*, 42–50. https://doi.org/10.1016/j.learninstruc.2014.09.003.

Lortie-Forgues, H., & Siegler, R. S. (2017). Conceptual knowledge of decimal arithmetic. *Journal of Educational Psychology*, *109*(3), 374–386. https://doi.org/10.1037/edu0000148.

Lortie-Forgues, H., Tian, J., & Siegler, R. S. (2015). Why is learning fraction and decimal arithmetic so difficult? *Developmental Review*, *38*, 201–221. https://doi.org/10.1016/j.dr.2015.07.008.

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Lawrence Erlbaum Associates. https://doi.org/10.5860/CHOICE.37-2294.

McCloskey, M. (2007). Quantitative literacy and developmental dyscalculias. In D. B. Berch, & M. M. Mazzocco (Eds.), *Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilities* (pp. 415–429). Paul H. Brookes Publishing.

McNeil, N. M., Fyfe, E. R., & Dunwiddie, A. E. (2015). Arithmetic practice can be modified to promote understanding of mathematical equivalence. *Journal of Educational Psychology*, *107*(2), 423–436. https://doi.org/10.1037/a0037687.

Moss, J. (1997). *Developing children's rational number sense: A new approach and an experimental program*. University of Toronto. https://doi.org/10.16953/deusbed.74839.

Moss, J., & Case, R. (1999). Developing children's understanding of the rational numbers: A new model and an experimental curriculum. *Journal for Research in Mathematics Education*, *30*(2), 122. https://doi.org/10.2307/749607.

National Governors Association Center for Best Practices, C. of C. S. S. O. (2010). *Common core state standards for mathematics*. National Governors Association Center for Best Practices, Council of Chief State School Officers. https://doi.org/10.21890/ijres.61754.

Ni, Y., & Zhou, Y. D. (2005). Teaching and learning fraction and rational numbers: The origins and implications of whole number bias. *Educational Psychologist*, *40*(1), 27–52. https://doi.org/10.1207/s15326985ep4001_3.

Rapp, M., Bassok, M., DeWolf, M., & Holyoak, K. J. (2015). Modeling discrete and continuous entities with fractions and decimals. *Journal of Experimental Psychology: Applied*, *21*(1), 47–56. https://doi.org/10.1037/xap0000036.

Ren, K., & Gunderson, E. A. (2019). Malleability of whole-number and fraction biases in decimal comparison. *Developmental Psychology*, *55*(11), 2263–2274. https://doi.org/10.1037/dev0000797.

Sformo, T. (2008). *Practical problems in mathematics: For automotive technicians*. Cengage Learning.

Shrager, J., & Siegler, R. S. (1998). SCADS: A model of Children's strategy choices and strategy discoveries. *Psychological Science*, *9*(5), 405–410. https://doi.org/10.1111/1467-9280.00076.

Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. Oxford University Press.

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., et al. (2012). Early predictors of high school mathematics achievement. *Psychological Science*, *23*(7), 691–697. https://doi.org/10.1177/0956797612440101.

Siegler, R. S., Im, S., Schiller, L. K., Tian, J., & Braithwaite, D. W. (2020). The sleep of reason produces monsters: How and when biased input shapes mathematics learning. *Annual Review of Developmental Psychology*, *2*(1). https://doi.org/10.1146/annurev-devpsych-041620-031544.

Siegler, R. S., & Lortie-Forgues, H. (2015). Conceptual knowledge of fraction arithmetic. *Journal of Educational Psychology*, *107*(3), 909–918. https://doi.org/10.1037/edu0000025.

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science*, *14*(3), 237–243. https://doi.org/10.1111/1467-9280.02438.

Siegler, R. S., & Pyke, A. A. (2013). Developmental and individual differences in understanding of fractions. *Developmental Psychology*, *49*(10), 1994–2004. https://doi.org/10.1037/a0031200.

Siegler, R. S., & Ramani, G. B. (2009). Playing linear number board games—But not circular ones—Improves low-income preschoolers' numerical understanding. *Journal of Educational Psychology*, *101*(3), 545–560. https://doi.org/10.1037/a0014239.

Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology*, *62*(4), 273–296. https://doi.org/10.1016/j.cogpsych.2011.03.001.

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science*, *10*(1), 89–96. https://doi.org/10.1111/j.1467-7687.2007.00569.x.

Tian, J. (2018). *Understanding percentages* (issue may). Carnegie Mellon University.

Tian, J., Braithwaite, D. W., & Siegler, R. S. (2020a). How do people choose among rational number notations? *Cognitive Psychology*, *123*, 101333.

Tian, J., Braithwaite, D. W., & Siegler, R. S. (2020b). Distributions of textbook problems predict student learning: Data from decimal arithmetic. *Journal of Educational Psychology*, *123*, 516–529. https://doi.org/10.1037/edu0000618.

Tian, J., & Siegler, R. S. (2018). Which type of rational numbers should students learn first? *Educational Psychology Review*, *30*(2), 351–372. https://doi.org/10.1007/s10648-017-9417-3.

Torbeyns, J., Schneider, M., Xin, Z., & Siegler, R. S. (2015). Bridging the gap: Fraction understanding is central to mathematics achievement in students from three different continents. *Learning and Instruction*, *37*, 5–13. https://doi.org/10.1016/j.learninstruc.2014.03.002.

Tyumeneva, Y. A., Larina, G., Alexandrova, E., DeWolf, M., Bassok, M., & Holyoak, K. J. (2018). Semantic alignment across whole-number arithmetic and rational numbers: Evidence from a Russian perspective. *Thinking and Reasoning*, *24*(2), 198–220. https://doi.org/10.1080/13546783.2017.1374307.

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP). (1990-2017). *Math assessment*. Retrieved from https://nces.ed.gov/NationsReportCard/nqt/Search/SearchOptions.

Varma, S., & Karl, S. R. (2013). Understanding decimal proportions: Discrete representations, parallel access, and privileged processing of zero. *Cognitive Psychology*, *66*(3), 283–301. https://doi.org/10.1016/j.cogpsych.2013.01.002.

Wynn, K. (2002). Do infants have numerical expectations or just perceptual preferences? *Developmental Science*, *5*(2), 207–209. https://doi.org/10.1111/1467-7687.00221_3.

This page intentionally left blank

CHAPTER TWO

Calibration and recalibration of stress response systems across development: Implications for mental and physical health

Megan R. Gunnar* **and Mariann A. Howland***
Institute of Child Development, University of Minnesota-Twin Cities, Minneapolis, MN, United States
*Corresponding authors: e-mail address: gunnar@umn.edu; howla042@umn.edu

Contents

Advances in Child Development and Behavior, Volume 63
ISSN 0065-2407
https://doi.org/10.1016/bs.acdb.2022.03.001

Abstract

Decades of human and animal research demonstrates that stress responsive neuroendocrine systems calibrate to the harshness of environmental conditions during fetal and early postnatal life. Emerging evidence indicates that if conditions change markedly over childhood, the hypothalamic pituitary adrenal (HPA) axis may recalibrate during puberty, another period that involves heightened neural plasticity and rapid maturation of neurobehavioral systems. These recent findings have prompted increased interest in the potential for stress system calibration/recalibration over development. To direct research in this area, this chapter integrates and discusses theoretical perspectives and empirical evidence pertaining to calibration and recalibration of the stress response. We describe how these concepts relate to other constructs, including sensitive periods, plasticity, and programming. We then consider four potential periods of calibration/recalibration: fetal, infancy, puberty, and pregnancy/lactation. In each section, we discuss evidence that the HPA and/or sympathetic medullary adrenal (SAM) system undergoes developmental change, rendering it more plastic and amenable to shift its activity in response to environmental conditions. We also review findings that the impacts of environmental harshness on stress responding persist beyond these periods. We then articulate that marked change in the quality of the environment (from harsh to benign or vice versa) is required in order for recalibration to occur, and that recalibration would result in shifts in stress responding to more closely align with the profiles of individuals who have experienced these conditions throughout life. Finally, we reflect on whether recalibration of the HPA and SAM system may extend to the other stress-responsive neurobehavioral systems.

1. Introduction

Environments vary in how much they support survival. In harsh environments, survival is difficult and requires significant tradeoffs to preserve reproductive fitness. In supportive environments, energy can be used more liberally without demands for limiting one aspect of functioning to preserve another (Del Giudice, Ellis, & Shirtcliff, 2011). Throughout evolution, humans have survived in a wide range of environments. Regulation of stress-system responding is critical to survival and adaptation (McEwen, 2017). Stable conditions allow for effective stress regulation; however, environments can change. Therefore, adaptive adjustment of response systems is necessary. There is a tension, nonetheless, between highly plastic systems that rapidly adjust to changing conditions and less plastic ones that are calibrated earlier and maintain their level of activity even when conditions change (see Reh et al., 2020). The study of critical periods in sensory development reflects this tension, revealing that there is an evolutionary advantage to establishing neural networks and then limiting their capacity to change. However, there is also a significant advantage in recalibrating stress-response systems when conditions change markedly.

Although stressors impact nearly all bodily systems (see Engel & Gunnar, 2020), this chapter will focus on the two primary stress-mediating neuroendocrine systems, the hypothalamic-pituitary-adrenocortical (HPA) and the sympathetic-adrenomedullary (SAM) systems. The neuroanatomy and physiology of these systems have been described elsewhere (e.g., Engel & Gunnar, 2020; Godoy, Rossignoli, Delfino-Pereira, Garcia-Cairasco, & de Lima Umeoka, 2018) and will not be covered here. Both of these systems have basal activity, circadian rhythms, and increased activity in response to stressors. Existing hypotheses about calibration and recalibration relate to the stress response (e.g., Del Giudice et al., 2011). Nevertheless, because stress responding impacts other aspects of system functioning, basal levels and the diurnal rhythm will also be considered.

The review begins by considering how the concepts of calibration and recalibration are related to other constructs, including programming. Then, four potential periods of calibration/recalibration are examined: fetal, infancy, puberty, and pregnancy/lactation. This last period focuses on women, although there might be a recalibration of stress responding in the pregnant woman's expecting partner, as well. In each section, the following criteria for calibration/recalibration are considered: (1) evidence that the HPA and/or SAM system are undergoing developmental change, and thus are more plastic; (2) evidence that during this period the system shifts in responding in relation to environmental harshness; (3) evidence that these effects on stress responding are preserved; and (4) evidence that stress-responding shifts with marked changes in environmental harshness (harsh to benign or the reverse). In each section, we consider whether changes in the activity of the HPA and SAM systems are associated with changes in behavioral responses to threat, and/or life history strategies (see the Adaptive Calibration Model; Del Giudice et al., 2011).

2. Theories and related constructs

To calibrate means to adjust precisely for a particular function. In the case of stress-responsive systems, calibration implies that the magnitude and duration of the stress response is adapted to the expected harshness of the environment. The term calibration in relation to stress-system functioning appears most prominently in the Adaptive Calibration Model (ACM) of Stress Responding (Del Giudice et al., 2011). However, evidence that early life events calibrate life-long patterns of stress responsiveness and emotionality predates the ACM (see Engel & Gunnar, 2020).

2.1 Adaptive calibration model

The ACM places evidence of early life calibration of the stress response system within a much larger, evolutionary-developmental context. According to the ACM, the stress response system is open to shaping by information about the harshness versus supportiveness of the environment, and, in turn, the stress system plays a central role in influencing life-history (LH) strategies. In low resource environments lacking the bioenergetic resources to support growth and reproduction, individuals shift to a slower LH strategy with reduced height, later puberty, later first birth, and fewer offspring. Although not discussed by Del Giudice et al. (2011), because activation of the stress response system is intensely energy demanding, one would expect a degree of hypo-responsivity to follow from such conditions. According to the ACM, once bioenergetic resources are sufficient, the key signals that calibrate the stress response system convey information about morbidity/mortality risk and the predictability of that risk. In childhood, these signals are conveyed through parental care, the predictability and controllability of events, threat stimuli, and social feedback. Rather than resulting in slower than normal LH strategies, unpredictable conditions with high signals of morbidity/mortality shift individuals towards fast LH strategies.

The ACM argues that LH strategies are particularly open to revision during different life stages, and it is calibration and recalibration of the stress response system that facilitates adjustments of these LH strategies. The ACM proposes that periods when the stress response is likely to calibrate and recalibrate are: prenatal stage (conception to birth), infancy (birth to about 2 years), childhood (3–6 years), middle childhood (7–11 years), and adolescence (12–17 years). In other words, as the child transitions through these life stages, changes in stress responsivity are expected if conditions of threat, quality of parental care, and predictability of these qualities of the environment have changed.

While the ACM proposes that the stress system is calibrated and recalibrated throughout childhood and adolescence, existing evidence from animal and human literature indicates that calibration/recalibration periods may be more limited. Experiments beginning in the late 1950's revealed that in rats and other altricial animals, maternal care in infancy shapes the HPA axis reactivity in ways that persist throughout the life of the animal. Furthermore, the same variation in maternal care that shapes the HPA axis response also influences fearfulness, response to novelty, risk-taking, and for females, the care they will provide their own offspring (see O'Donnell & Meaney, 2020). Work by Michael Meaney and colleagues traced many of these effects to changes in the epigenome (see O'Donnell & Meaney, 2020). Initially, work focused on methylation of specific genes. For the HPA axis, the glucocorticoid receptor (GR) gene in limbic and forebrain regions was shown to be more or less methylated by maternal care. More recently, attention has shifted towards whole genome studies of DNA methylation. In a recent human study of a large subsample of children from the AVON study, an early childhood sensitive period model fit the data better than a cumulative adversity or recent adversity model (Dunn et al., 2019). These and other findings speak to an important outstanding question: is the stress system calibrated and recalibrated fairly continuously throughout childhood and adolescence (as the ACM would suggest), or is there a more limited set of sensitive periods for stress systems calibration/recalibration?

2.2 Sensitive periods

Sensitive or critical periods have been studied extensively for brain circuits involved in sensory and linguistic functions. During sensitive periods for specific circuits, those circuits become more open to experience. The mechanisms opening and closing these sensitive periods involve parvalbumin cells, a shift in excitatory-inhibitory balance that opens the period, and the development of perineuronal nets that close the period and "hold" the circuit in place (Reh et al., 2020). The term sensitive, as opposed to critical, indicates that while a system may exhibit heightened sensitivity during a particular developmental period, it can be altered at earlier or later points, just not as easily. While we do not know what opens and closes sensitive periods for the HPA and SAM system, similar mechanisms may be involved. For example, there is recent evidence that thyroid hormones (TH) play a role in sensitive period regulation (Batista & Hensch, 2019), and it is well established that TH is stimulated by maternal licking in the rat and triggers a cascade of events resulting in increased transcription of the GR in the hippocampus which regulates HPA axis responsivity (O'Donnell & Meaney, 2020).

The idea that stress responses are calibrated and recalibrated during sensitive periods fits with the fact that both the SAM and HPA axis are open to activity-dependent alterations in their reactivity and regulation under conditions of intense, prolonged, or chronic stress throughout life. These activity-dependent alterations in stress responding are discussed in the context of allostasis, or the process of maintaining stability (i.e., homeostasis) through change, through activation of stress hormones and other stress mediators (McEwen, 2017). Activating stress responses frequently produces allostatic load or overload, which refers to wear and tear on the body and brain. It also dysregulates stress-response systems, producing hypo-responding (not turning on adequately when needed) or hyper-responding (not turning off when the stress or threat is passed). Furthermore, as they are dynamic and constantly changing, allostatic adjustments to system set points tend to occur gradually, and once a prolonged period of exposure to stressors is over, there may be a return to activity that resembles pre-challenge dynamics. In contrast, the calibration of stress responding that occurs during sensitive periods has impacts on the system that persist over prolonged periods of development.

2.3 Fetal programming, DOHaD and predictive adaptive response

The fetal programming ("Barker") and Developmental Origins of Health and Disease (DOHaD) hypotheses are also relevant to calibration of the stress response (Vickers, 2022). As originally conceived, these hypotheses pertain to signals (e.g., nutrient, stress) of the quality of the ex-utero environment which shape fetal growth and program set points of body systems, with the potential to confer risk for later disease. A related evolutionary-developmental perspective, the Predictive Adaptive Response hypothesis, casts these programming effects as adaptations that prepare the fetus for survival in the extra-uterine environment (Bateson, Gluckman, & Hanson, 2014). If nutrition is poor and stress signals are high, the fetus develops a more reactive stress response and a "thrifty phenotype," or the capacity to extract all the nutrition it can from what is available and to conserve energy at the expense of linear growth. It may also speed up its development so that it can survive if born prematurely (see Sandman, 2018). Health problems arise if this prediction does not match the extra-uterine environment, that is, when postnatal conditions are benign and nutrient rich. In this case, rapid weight gain, overweight and obesity, abdominal adiposity, increased inflammation, early onset cardiometabolic disease, and premature death are increased in likelihood. Evidence that a mismatch between pre- and postnatal environments increases later physical and mental health risks, of course, is inconsistent with the idea that the stress response system recalibrates during different periods of development. Even so, some argue that a stressful prenatal environment actually increases sensitivity to the postnatal environment and thus the opportunity for recalibration in benign conditions (Pluess & Belsky, 2011).

2.4 Summary

The concept of stress system calibration transects a number of literatures. It has a long history in the field of developmental psychobiology, with the earliest evidence showing that the HPA axis is shaped or calibrated by maternal care. Calibration of the stress response system is central to the Adaptive Calibration Model, though this model argues for nearly continuous calibration from the prenatal period through adolescence, notably more at transitions across age periods. However, research beginning with the study of early postnatal experience in the rat and now extended to humans has tended to support sensitive periods in early development, likely before age 2 in humans, with perhaps another sensitive period around puberty. This would align calibration and recalibration with sensitive periods in neural circuit development, although it is unclear the extent to which mechanisms that open and close sensitive periods in sensory system development also contribute to the opening and closing of such periods for stress responding. Finally, DOHaD research and work on sensitive periods in the calibration of the stress system also overlap and this highlights both the importance of prenatal development, but also the confluence of nutrition and stress.

3. Fetal period

While existing theory and research on early life calibration of the stress response has primarily focused on the postnatal period, this process is increasingly understood to begin during fetal life. During gestation, the brain and other organ systems are rapidly developing and highly plastic. Most of the brain's billions of neurons are formed by mid-gestation, and the fetal brain connectome exhibits adultlike network properties (Thomason, 2020). Several decades of evidence indicates that the fetal stress response system is programmed or calibrated to match ex-utero environmental conditions. A meta-analysis across 14 vertebrate species demonstrates that in utero programming of the HPA axis is an evolutionarily ancient, highly conserved mechanism, consistent with the notion that early calibration of the stress system shapes life history behaviors and health outcomes (Thayer, Wilson, Kim, & Jaeggi, 2018).

3.1 Prenatal development of the stress response system

The basic architecture of the stress response system is established in the fetal period. The SAM system begins to develop in the first trimester, when medullar cells arise from the neural crest and migrate to invade the adrenal and form primitive sympathetic ganglia (Wells, 2015). Chromaffin endocrine cells are apparent in the medulla by around 10 weeks' gestation, though at birth, there are still extra-adrenal chromaffin cells that atrophy in the months following. Regarding the HPA axis, its developmental timetable has been described in detail elsewhere (see Howland, Sandman, & Glynn, 2017). One seminal study showed that by mid-to-late gestation, the fetus can mount a cortisol response partially independent of the mother's (Gitau, Fisk, Teixeira, Cameron, & Glover, 2001). However, the coordinated release of hormones from the human fetal HPA does not appear to commence until late gestation, when the fetal adrenal cortex resembles a rudimentary form of the adult adrenal. Because the stress system is immature in fetal life, it is heavily shaped by maternal and placental inputs.

3.2 Potential mechanisms of stress-system calibration

Environmental stressors (e.g., natural disaster, famine) are associated with persisting differences in HPA axis responsivity, in the structure and function of brain regions regulating the stress response, and in risk for stress-related mental and physical disorders (Szutorisz & Hurd, 2016; Yong Ping et al., 2020). Animal models demonstrate the causal role of harsh prenatal conditions in inducing long-term alterations in offspring HPA axis activity via epigenetic mechanisms (Cao-Lei et al., 2020; Creutzberg et al., 2021). Importantly, the fetus also adjusts its developmental trajectory in response to supportive conditions (e.g., Katzow, Messito, Mendelsohn, Scott, & Gross, 2019), reflecting that plasticity confers both vulnerability and opportunity in development. In utero programming or calibration of the stress response is thought to occur through maternal and placental nutritional, vascular, immune, and endocrine signals which carry information about environmental conditions.

Among these signals and especially relevant to the developing HPA axis are placental corticotropin releasing hormone (CRH) and maternal cortisol. Placental CRH (pCRH) is essentially identical to hypothalamic CRH (reviewed in Sandman, 2018). Levels increase exponentially over gestation in both the fetal and maternal compartments. pCRH is involved in regulating the timing of labor and delivery, with accelerated pCRH trajectories predictive of premature birth. Because it is responsive to other stress mediators (e.g., nutritional, immune, vascular), pCRH is proposed to be an integrative signal of the quality of the anticipated postnatal environment. Several investigations also demonstrate associations between pCRH and psychosocial stressors experienced by the mother during pregnancy and her own childhood (see Steine et al., 2020).

pCRH may calibrate the fetal stress response by stimulating fetal production of cortisol and/or by increasing responsiveness of the fetal adrenal to ACTH (see Howland et al., 2017). pCRH presumably crosses the immature fetal blood-brain-barrier, and CRH receptors are widely expressed in the brain from 13 weeks' gestation, particularly in regions regulating stress responses. A rich rodent literature demonstrates effects of CRH on the immature brain. In terms of the development of key limbic regions (e.g., hippocampus), the first week of life in the rodent is roughly equivalent to the third trimester in humans (Avishai-Eliner, Brunson, Sandman, & Baram, 2002), and CRH exposures during this period reduces dendritic branching in hippocampal and cortical neurons (Curran, Sandman, Davis, Glynn, & Baram, 2017). In humans, prospective longitudinal studies suggest that elevated pCRH is associated with cortical thinning and elevated psychiatric symptoms in childhood (see Sandman, 2018).

Exposure to elevated maternal cortisol may calibrate the fetal HPA axis directly or via modulation of placental function (see Krontira, Cruceanu, & Binder, 2020). Maternal cortisol levels increase 2- to 5-fold over gestation because of a positive feedback loop between maternal cortisol and pCRH. Placental production of the enzyme 11β-hydroxysteroid dehydrogenase 2 (11β-HSD-2) limits fetal exposure to maternal cortisol by oxidizing cortisol into more inert cortisone; however, an estimated 15% of maternal cortisol does cross the placenta. By mid-to-late gestation, maternal and fetal cortisol are significantly correlated. *11β*-HSD-2 production is lower during early and late gestation, suggesting that the fetal HPA axis may be more susceptible to maternal cortisol during these periods. Interestingly, several stress mediators are shown to downregulate placental 11β-HSD-2 production, including catecholamines, cytokines, and hypoxic factors (see Howland et al., 2017). Thus, stress may increase the vulnerability of the fetus to maternal cortisol.

There are multiple avenues through which exposure to elevated maternal cortisol might program the fetal HPA axis. One path would be through altering glucocorticoid (GR) and mineralocorticoid receptor (MR) density in the brain, pituitary, and adrenal. GRs and MRs are highly expressed in the developing brain and in most tissues by mid-gestation (Diaz, Brown, & Seckl, 1998; Noorlander, De Graan, Middeldorp, Van Beers, & Visser, 2006). In animals, prenatal synthetic glucocorticoid exposure increases CRH mRNA, GR, and MR expression in the amygdala (see Welberg & Seckl, 2001) and decreases GR and MR expression in the hypothalamus, hippocampus, and PFC (Bingham, Rani, Frazer, Strong, & Morilak, 2013; Xu et al., 2018), resulting in increased anxiety and depression-like behavior.

3.3 Evidence of long-term effects of fetal stress exposure

Prenatal elevations in cortisol are associated with persisting alterations in HPA axis activity. Children born at term and exposed to synthetic glucocorticoids prenatally (standard of care for pregnant women at risk for preterm delivery) demonstrate increased cortisol responses to painful stressors as neonates (Davis, Waffarn, & Sandman, 2011) and to the Trier Social Stress Test for Children (TSST-C) at age 6–11 (Alexander et al., 2012). Elevations in maternal cortisol are associated with larger responses to pain stressors in neonates and infants (Davis, Glynn, Waffarn, & Sandman, 2011) (Irwin et al., 2021) and to the TSST-C in 6-year-old children (Simons, Zijlmans, Cillessen, & de Weerth, 2019; though some studies report blunted responses; e.g., Nazzari et al., 2019). Prenatal exogenous and maternal glucocorticoid exposures also relate to brain structure and function (e.g., Buss et al., 2012; Graham et al., 2019) and risk for mental and physical health problems in childhood (see Zijlmans, Riksen-Walraven, & de Weerth, 2015).

Because of obvious challenges in assessing the fetal stress system, direct evidence of in utero calibration is limited. Prenatal undernutrition and glucocorticoid exposures are shown to confer risk for low birthweight, a long-regarded indicator of prenatal adversity, and low birthweight is associated with postnatal stress reactivity (Osterholm, Hostinar, & Gunnar, 2012; Reynolds, 2013). Other studies have associated prenatal stressors with postnatal HPA axis functioning, from the neonatal period into adulthood (see Howland et al., 2017). These findings are naturally limited by the potential for confounding by postnatal influences, though most investigations attempt to account for postnatal factors. Another limitation is that studies have mostly relied on indicators of maternal *distress* (e.g., depression, anxiety) rather than more objective measures of environmental conditions, yet these indicators are typically only modestly correlated.

Quasi-experimental disaster studies offer some of the best evidence in humans stress system is calibrated in utero, with persisting influences on offspring stress-related behavioral outcomes. These studies allow for examination of the timing of effects during fetal development, and, given the time-limited nature of disasters, may also offer an opportunity to examine whether recalibration occurs later in development if conditions have changed. Results from Project Ice Storm, a cohort of women who were pregnant during the 1998 Quebec ice storm and their children, link greater objective hardship (e.g., loss, threat) from the storm to heightened cortisol reactivity to the TSST-C at age 13 (Yong Ping et al., 2020). Similar results were found in women exposed to the 2008 Iowa floods during pregnancy and their 2-year-old children using a separation stressor (Yong Ping et al., 2015). Findings also extend to basal activity of the axis (Nomura et al., 2021).

Tests of potential mechanisms in humans are inconclusive, perhaps because the pathways are numerous and complex. In addition to some support for the roles of maternal cortisol and placental CRH, there is evidence of changes in the methylation of HPA axis-relevant genes (e.g., genes encoding 11β-HSD-2, CRH-binding protein) following prenatal exposure to natural disaster or war (Kertes et al., 2016; Nomura et al., 2021). To advance understanding in this area, more research measuring other stress mediators is needed (e.g., immune, nutritional), along with designs that would allow statistical examination of mediating pathways.

3.4 Summary

The fetal period involves rapid growth and development, and the architecture of the stress system is fundamentally established during this time. Consistent with the concept of calibration, harsh ex-utero environments are associated with persisting differences in offspring HPA axis activity, a process presumably mediated by alterations in maternal and placental biology. The strongest evidence of HPA axis calibration during fetal life comes from quasi-experimental natural or manmade disaster studies, taken together with findings from experimental animal models. However, the stress system is far from fully mature at birth, and its development continues to be strongly regulated by inputs from caregivers into the postnatal period.

4. Infancy

In animal models, there is good evidence that prenatal stress effects can be partially reversed through postnatal experience (e.g., George, Stout, Tan, Knox, & Liberzon, 2013), raising the likelihood that if there is a sensitive period for calibrating the stress response system, it extends into the postnatal period.

4.1 Postnatal development

Sympathetic tone is high in infancy, decreasing across development, while parasympathetic activity is low, increasing rapidly over infancy and childhood (Harteveld et al., 2021). The HPA axis undergoes a number of changes during infancy (see Engel & Gunnar, 2020). At birth, the adrenal cortex includes a large fetal zone that will involute over the first 6 months postnatal. Its remaining zones are present but not organized as they will be at maturity. The immature neonatal liver produces relatively low levels of corticosteroid binding globulin, so most circulating cortisol is unbound and biologically active. Total cortisol in circulation is low, although the unbound portion is still probably higher in early infancy than later in life. Binding globulin levels increase over the first few months, resulting in higher total cortisol output, but a reduction in free hormone levels. Fast feedback mechanisms mature in early infancy, resulting in a decrease in the magnitude of cortisol response to mildly arousing stimulation (e.g., undressing and weighing), while the capacity to respond to more intense stimulation (e.g., inoculations) is retained. The mature circadian rhythm in cortisol is not present at birth. By 6 weeks a single peak aligned with time-of-day emerges, while the fully adult-like rhythm is not observed until the child gives up the afternoon nap. Using area under the curve from ground (AUCg), when cortisol is assessed across the day, the AUCg is higher among 12,-18- and 24-month-olds than among 30- and 36-month-old children (Watamura, Donzella, Kertes, & Gunnar, 2004).

4.2 Differential activity during infancy

There is good evidence that the infant HPA axis responds differently as a function of the quality of stimulation. Exposure to pain stressors sensitizes the system, leading to larger responses to subsequent painful stimuli in neonates (Gunnar, 1992). Within the first months, insensitive care during bathing slows the cortisol return to baseline (reviewed in Engel & Gunnar, 2020). By a year of age, infants in higher risk neighborhoods with more household chaos produce more cortisol over time (Tarullo, Tuladhar, Kao, Drury, & Meyer, 2020). With the formation of the attachment relationship, the presence of the attachment figure in secure relationships provides a powerful buffer of the stress response (discussed in Engel & Gunnar, 2020). For example, poverty is associated with elevated cortisol in response to the challenge of a well-child visit with inoculations, but only in insecure attachment relationships. Thus there is good evidence that the harshness versus supportiveness of the environment in infancy affects concurrent stress responding. The question is whether these effects program the axis in ways that carry forward into later periods of development.

4.3 Postnatal calibration

The best evidence for postnatal calibration comes from studies of infants reared in institutions (e.g., orphanages) until they are placed in families. These children experience marked changes in care from care that typically fails to support normative development to care that allows considerable recovery from earlier deprivation (reviewed in Gunnar & Reid, 2019). Studied while still in institutional care, there is a dysregulation of the normal diurnal rhythm, with low levels being observed early in the morning when peak levels are expected. A relatively flat pattern of cortisol production across the day is found, which is more profound among preschool-aged children who experienced poorer social care in their institution. Critically, this pattern continues for a number of years after placement in supportive, family care.

Consistent with the argument that harsh environments that do not have sufficient bioenergetic resources would lead to down-regulation of the HPA stress response (Del Giudice et al., 2011), children in institutional care and those adopted or fostered out of the institutions exhibit stress hyporesponsiveness (see Gunnar & Reid, 2019). They also show linear growth delay, often severe, and rapid catch-up growth once placed in families. Whether early adverse care produces hypo- or hyper-responsiveness of stress systems, however, is unclear. Some argue that deprivation and threat produce different sequelae (McLaughlin & Sheridan, 2016), and there is some evidence that in adolescence this might be true (Doom et al., 2020). However, whether deprivation versus threat in infancy differentially shapes stress hyper- versus hypo-responsiveness is still an open question.

Studies of post-institutionalized (PI) youth suggest that sometime in the latter half of the second year of the life, the window of calibration of the stress response begins to close (Dahl & Gunnar, 2009). This has been shown in the Bucharest Early Intervention Project (BEIP). The BEIP randomly assigned children to remain in care-as-usual or to be removed from institutional care and placed in research-supported foster care (McLaughlin et al., 2015). The children ranged in age at time of randomization, thus allowing a test of a sensitive period for calibrating the stress response system. When tested at 12 years of age in the TSST-C, children placed in families before

24 months of age showed cortisol and pre-ejection period (PEP) responses comparable to never institutionalized, comparison children. Children placed at or over 24 months showed responses that were as blunted as those of the care-as-usual children (McLaughlin et al., 2015).

Even variations in care within non-maltreating ranges may calibrate the HPA axis in infancy. Center-based care in early childhood is associated with a rise, rather than the expected diurnal decline, in cortisol (Vermeer & van IJzendoorn, 2006), and maternal insensitivity is associated with insecure attachment (De Wolff, & van IJzendoorn, M.H., 1997) and, thus, poorer buffering of the stress system (for review, see Engel & Gunnar, 2020). In one study, center-based care and higher maternal insensitivity before age 3 was associated with a lower wake up cortisol level at age 15 and thus likely a flatter diurnal rhythm (Roisman et al., 2009). In another longitudinal study, at age 37, cortisol reactivity to the TSST was larger with higher stressful life events in infancy and early childhood (Young et al., 2021).

4.4 Summary

All of the studies examined in this section suggest that conditions in the first years following birth influence the long-term functioning of the HPA axis and perhaps the SAM system as well. Although the clearest causal evidence comes from the BEIP study, other longitudinal studies of less harsh early conditions also predict HPA axis activity years following infancy. We now turn to whether there are periods of recalibration post-infancy.

5. Adolescence and puberty

Adolescence is a period of dramatic change that transitions the individual from childhood dependence on caregivers to adulthood independence (reviewed in Dahl & Gunnar, 2009). It is a time of heightened neural plasticity and active synaptic pruning. It is also a time when the neurocircuitry underlying stress reactivity and regulation undergoes its final maturation. While humans are at the peak of their physical health in adolescence, they are more vulnerable to stress-related disorders, including depression and anxiety (Dahl & Gunnar, 2009). The puberty-HPA increase-psychopathology risk hypothesis proposes that increased stress responsivity underlies this increased psychiatric risk (e.g., Romeo, 2018; Spear, 2000). Plasticity, however, cuts both ways, and if the stress system is in a plastic state as it finishes maturing, then it may well also recalibrate in response to supportive conditions.

5.1 Adolescence versus puberty

Adolescence is defined by age, beginning roughly around age 10 and finishing around age 19, although some have used other divisions, considering middle childhood to last until age 12 and adolescence to encompass the "teen" years (i.e., 13–19). In contrast, puberty is the process of physical change that matures the child's body into an adult, reproductively capable body. Pubertal development beings with adrenarche at 6–7 years which is marked by the gradual increase in the adrenal androgen dehydroepiandrosterone (DHEA). DHEA and its sulfated version, DHEAS, are neurosteroids, and their rise from age 6 through the early 20s is believed to play a role in extended brain maturation in humans (reviewed in Quinn, Greaves, Badoer, & Walker, 2018). During adrenarche, DHEA may act on the amygdala to reduce fearfulness, increasing the range of social interaction with unfamiliar peers, and on the hippocampus, promoting memory and reduced vulnerability to stressors encountered as the child moves out of the protective parental sphere. It may also facilitate brain rewiring in response to changing social environments.

Gonadarche begins with rises in gonadotrophins secreted by the hypothalamus followed by increases in sex steroid production by the gonads. Gonadal hormones stimulate libido and the development of secondary sexual characteristics. Gonadarche happens in girls at roughly 10–11 years and in boys at 11–12 years and in both sexes is typically complete by 17–18 years. Pubertal development is correlated with rises in adrenal and gonadal steroids but is most often measured from external evidence of sexual maturation (e.g., five Tanner stages; Peper & Dahl, 2013).

5.2 Animal models

Much of the evidence for adolescence as a period of stress response recalibration comes from work in rats and mice (see Romeo, 2018). Animal models have the advantage that the timing of stress-eliciting experience can be controlled. In rats and mice, there is a marked maturation in stress responding from preadolescence (middle childhood in humans) to adulthood. Preadolescent rats show a protracted HPA stress response, while the adult animal can efficiently shut off the response. There is evidence that this shift occurs abruptly around mid-adolescence and involves changes in multiple brain regions that provide input to the CRH-producing hypothalamic cells, as well as changes within the axis itself. Stressors also produces changes in stress circuitry in the brain, with some evidence that effects are opposite in adolescent compared to adult rats. Stress in adults increases spine density in the amygdala and reduces it in the hippocampus and prefrontal cortex, but the opposite pattern of changes is noted in the adolescent rat (Eiland & Romeo, 2013; Romeo, 2017). Stress also produces different changes in gene expression in adolescent versus adult rats (Eiland & Romeo, 2013). Although it is argued that the stress system is highly conserved and thus animal models can be used to understand stress system development during adolescence in humans, there are challenges in translation. For example, it is not clear that the rodent has a period equivalent to adrenarche. Furthermore, the prediction from animal work would be that while adolescence may be a period of recalibration, it would also be one of *less* reactivity and more rapid return to baseline than observed in younger children. This is not what is often seen, as we review below.

5.3 Pubertal change in stress responding in humans

Although the puberty-HPA increase-psychopathology risk hypothesis (e.g., Spear, 2000) pertains to stress reactivity, researchers advocating it often point to evidence that basal activity of the HPA axis increases from childhood to adolescence (see Engel & Gunnar, 2020). While much of this evidence is cross-sectional, longitudinal data also support an increase in cortisol production with puberty, which in girls appears to occur around the mid-point (i.e., Tanner Stage III).

Some of the best evidence that puberty is a period of changing regulation of the HPA axis comes from studies examining how pre-pubertal life stress relates to HPA axis activity earlier and later in puberty. In one study, more severe early life stress was associated with blunting of the cortisol awakening response (CAR) for youth early in pubertal development, while for youth later in pubertal development it was associated with a heightened CAR (King et al., 2017). These authors also demonstrated longitudinally that both within- and between-participants, the coupling of DHEA and cortisol increased from weaker to stronger across pubertal development, though this pattern was attenuated for youth with more severe life stress histories (King, Graber, Colich, & Gotlib, 2020).

Regarding changes in stress reactivity, age rather than pubertal stage has been examined most often, with children being defined as those 7–12-years (roughly the period of adrenarche) and adolescents as 13–17 years (roughly the period of gonadarche). Using these age parameters, a recent meta-analysis of responses to social evaluative, public speaking stressors reported that children exhibit larger cortisol and autonomic responses than do adolescents (Seddon et al., 2020). While this would be consistent with the rodent models, the preponderance of the 57 studies examined either children or adolescents, with only a few made comparisons within the same study. Furthermore, many of the children and adolescents were from clinical or other high-risk populations which typically exhibit blunted responses.

A markedly different picture more consistent with the puberty-HPA increase-psychopathology risk hypothesis emerges when researchers compare children, adolescents and/or adults in the same study. Yim and colleagues conducted three studies (Yim, Quas, Cahill, & Hayakawa, 2010; Yim, Quas, Rush, Granger, & Skoluda, 2015), using the TSST. Two studies found no differences in stress reactivity between children and adults. Two studies compared children and adolescents, with one finding no difference and the other finding a larger stress response among adolescents than children. Gunnar and colleagues (Gunnar, Wewerka, Frenn, Long, & Griggs, 2009) assessed 9-, 11-, 13-, and 15-year-olds in the TSST and observed a larger cortisol response among the adolescents than children. In another study, adolescents responded more to the TSST and to a social rejection stressor than did children (Stroud et al., 2009). Finally, in the only study to employ a longitudinal analysis (van den Bos, de Rooij, Miers, Bokhorst, & Westenberg, 2014), 8–17-year-olds were tested in a public speaking task twice, 2 years apart. Both cortisol and sAA responses increased between- and within-participants, and these increases were more closely associated with pubertal development than age. One challenge in interpreting these human data is the use of social evaluative stressors. While social evaluative threat is the most effective laboratory stressor (Dickerson & Kemeny, 2004), adolescence is also a period of heightened sensitivity to social evaluation (e.g., Somerville et al., 2013). This makes it difficult to know whether there is a developmental change in stress reactivity with puberty, or whether the stressor used is more potent for adolescents than for children or adults.

What if we bypass the limbic system and stimulate the stress system pharmacologically? In the one study that did so (Stroud, Papandonatos, Williamson, & Dahl, 2004), participants ages 6–16 were carefully screened to be psychologically and physically healthy, and pubertal stage was determined through physical exam. Participants had an indwelling catheter inserted and adapted to this and the laboratory setting for 24 h before corticotropin-releasing hormone (CRH) was infused. Repeated plasma samples revealed subtle increases in the pituitary-adrenal response to

CRH with pubertal stage in girls but no changes for boys. Girls showed increased total cortisol responses with lower peak levels but a prolonged period of elevation prior to a return to baseline. There were no effects of puberty noted for boys. Since increased reactivity to social evaluative stressors is noted for both sexes, this suggests that either CRH-producing cells in the hypothalamus respond more to limbic inputs in adolescence, the developmental changes are extra-hypothalamic, or social evaluation is a more potent stimulus for adolescents than children or adults. These possibilities need to be sorted out.

5.4 Recalibration

Tests of recalibration require that conditions change markedly from one life period to the next. In humans, it is difficult to find clear tests of stress system recalibration because conditions tend to be correlated across life periods. Even when we can identify a discretely stressful event (e.g., natural disaster with clear onset and offset), individuals who are the most affected are likely to have both prior histories of adversity and a lack of resources to facilitate quick recovery (discussed in Gabard-Durnam & McLaughlin, 2019).

As with tests of calibration in infancy, previously-institutionalized (PI) youth provide the best tests of the pubertal stress recalibration. PI youth typically experience a marked change in life conditions typically long before puberty. Recalibration could be concluded if PI youth show differences in stress responding from low-adversity children prior to puberty, but then with puberty shift towards functioning comparable to low-adversity youth. This has been observed cross-sectionally for the CAR (Quevedo, Johnson, Loman, LaFavor, & Gunnar, 2012) and for cortisol stress responses to the TSST (DePasquale, Donzella, & Gunnar, 2019). More importantly, for the cortisol stress response, this pattern has also been observed longitudinally within-participants (Gunnar, DePasquale, Reid, Donzella, & Miller, 2019). This study employed an accelerated longitudinal design, with the youngest participants at the first assessment being 7- and 8-years-old and thus either pre-pubertal or in early adrenarche, and the oldest children at the first assessment being 13–14 years. Almost all the PI youth were adopted before age 3, and their families and those of the non-adopted, comparison children were highly resourced. Participants were assessed three times over 2 years. There were marked differences in cortisol responses to the TSST among the groups at the earliest stages of puberty, while, at the highest pubertal stages the groups did not differ. Recalibration involved a reversal of the hypo-cortisolism reported for young children both while in institutional care and in the years following adoption into supportive homes (reviewed in Gunnar & Reid, 2019). A similar recalibration result was obtained in a study of left-behind children in China who were reunited with their parents before puberty or not with their parents before puberty (Zhang et al., 2021). In addition, in two studies when DHEA was examined in the children exhibiting HPA stress response recalibration, this recalibration was also noted in the increased coupling of DHEA and cortisol over pubertal development in the PI youth to levels that were comparable to those noted throughout puberty in the comparison youth (Howland, Donzella, Miller, & Gunnar, 2020; King et al., 2020).

As discussed earlier, the BEIP study also used a modified TSST to measure stress-system responding among children who were institutionalized as infants and toddlers and then randomly assigned to foster care versus care-as-usual. When the children were 12-years-old, it was found that both those randomized to foster care at 24 months or later and the care-as-usual children had blunted HPA and SAM system responses compared to a community control group (McLaughlin et al., 2015). However, this finding was qualified by quality of care and stressful life events at age 12 (Wade et al., 2020). The community comparison youth showed evidence of blunted cortisol responding if parenting quality was poor at age 12, and for youth who had been randomly assigned to foster care (i.e., removal from adversity), no evidence of HPA axis recalibration was obtained (note that the children were likely still in adrenarche). However, SAM system activity did show evidence of recalibration. Specifically, for youth randomized to foster care, SAM system activity was not blunted and was similar to comparison children if stressful life events were low at age 12, but was blunted if stressful life conditions were high.

5.5 Recalibration and behavior

If the systems that respond to threat and stress are recalibrating with puberty when conditions shift from harsh to benign, then might behavior shift from being more fearful and anxious than comparison youth to similarly fearful and anxious? Certainly, there is evidence that with puberty, subcortical systems, including the amygdala, show evidence of increased plasticity and change (Laube, van den Bos, & Fandakova, 2020). However, as far as a pubertal recalibration of the threat-response neurobehavioral system, this is not what has been observed.

Again, focusing on PI youth, there is relatively little evidence of heightened fearfulness prior to puberty in these children reviewed in Gunnar & Bowen, 2021; Gunnar & Reid, 2019). Furthermore, measures of anxiety and depression tend to increase with the transition from childhood to adolescence and then adulthood for previously-institutionalized more so than for comparison youth, perhaps related to failures in accomplishing the developmental tasks of adolescence and early adulthood (Golm et al., 2020). Notably, we have recently found that as the HPA axis recalibrates, there is an increase in internalizing symptoms in PI youth (Perry, DePasquale, Donzella, & Gunnar, 2020; Perry, Donzella, & Gunnar, 2022). On the other hand, the BEIP project has recently presented intriguing findings suggesting that positive family experiences in adolescence may have a more profound effect on outcomes at age 16 years than similar experiences in middle childhood (Colich et al., 2021). Thus, how and whether stress hormone recalibration translates into recalibration of behavioral threat and emotion systems in adolescence remains an open question.

5.6 Summary

The pubertal period is one of high plasticity, and as neuroendocrine, brain and behavioral systems are rapidly maturing, they are likely sensitive to changes with experience. Animal models may provide some guide to our understanding of human development, but there are many ways in which the peripubertal period in humans is different than in these model animals. Evidence of recalibration of the stress-response system is the clearest in work on youth who were adopted or fostered as young children from harsh, institutional conditions to more benign and supportive family care. For these individuals it does appear that puberty, beginning with its earliest stages (i.e., adrenarche) begins to open a window for recalibration, which results in shifts from blunted stress responding to reactive profiles comparable to those of youth who were conceived and raised in more benign, supportive conditions.

6. Pregnancy and lactation

Whether potential recalibration periods exist beyond puberty remains to be determined. Pregnancy and lactation may be additional windows in the female lifespan. Pregnancy involves dramatic surges in both reproductive and adrenal hormones, which support gestation, birth, lactation, and caregiving behavior (see Almanza-Sepulveda, Fleming, & Jonas, 2020). Pregnancy is increasingly recognized as a sensitive period of heightened maternal brain plasticity (Glynn, Howland, & Fox, 2018), with animal models indicating this plasticity extends through lactation. These are two conditions that raise the possibility of stress system recalibration. Below we consider this information and examine whether evidence exists in support of the potential for recalibration across these periods.

6.1 Changes in stress responding during pregnancy and lactation

Normative pregnancy and lactation involve dramatic changes in the stress system. Concentrations of HPA axis hormones in maternal circulation shift markedly over pregnancy, with progressive rises in placental CRH, maternal ACTH, and maternal cortisol (see Howland et al., 2017). The maternal SAM system likely participates in this process (Petraglia, Sutton, & Vale, 1989). After birth, maternal cortisol levels gradually decline to non-pregnant levels as the hypertropic adrenals progressively downsize (Magiakou et al., 1996). A transient adrenal suppression results from low hypothalamic CRH secretion, which is proposed to increase women's vulnerability to mental health difficulties. Basal hormone levels are further altered in lactation (see Hasiec & Misztal, 2018), as breastfeeding bouts are associated with acute reductions in ACTH and cortisol (Handlin et al., 2009).

Pregnancy and lactation also involve adaptations to the stress response. Over gestation, HPA axis responsivity is progressively attenuated, which is thought to protect the maternal-fetal dyad from the damaging effects of excessive stress system activation (Slattery & Neumann, 2008). By late pregnancy, exogenous CRH does not induce an ACTH or cortisol response, and physical and psychosocial laboratory stressors either produce a dampened or lack of response (see de Weerth & Buitelaar, 2005; Slattery & Neumann, 2008). Women who do not show this adaptive downregulation appear to be at increased risk for shortened gestation (Glynn, Dunkel Schetter, Hobel, & Sandman, 2008) and lower infant birth weight (Kivlighan, DiPietro, Costigan, & Laudenslager, 2008). Interpreted from the perspective of calibration, the degree of attenuation in stress reactivity over gestation may reflect adaptive calibration of the maternal (and fetal) stress response to environmental harshness. Blunting of the stress response continues in lactation (e.g., Cox et al., 2015; Magiakou et al., 1996).

6.2 Maternal brain plasticity

Heightened neural plasticity supports the development of the "maternal brain circuit," which includes stress system neurocircuitry (Leuner & Sabihi, 2016). An extensive rodent literature demonstrates that this plasticity exerts lasting impacts on maternal brain and behavior, with parallel evidence in humans emerging (see Kim, 2021). In rodents, the pregnant and lactating brain is altered by many of the same neural processes involved in forming neural circuits during early development, with neural, dendritic, and synaptic plasticity evident in numerous brain regions, including those regulating the stress response (see Leuner & Sabihi, 2016; Slattery & Hillerer, 2016). This brain remodeling is critical for caregiving behavior but also exerts broader influences, with effects on learning, memory, and anxiety-like behaviors that persist into old age (e.g., Gatewood et al., 2005; Love et al., 2005).

In human mothers, first pregnancy involves marked reductions in gray matter volume observable until at least 2 years post-pregnancy which are positively associated with self-reported attachment and neural responsivity to one's own infant (Hoekzema et al., 2017). These reductions are comparable to those seen over puberty (Carmona et al., 2019). Several studies document increases in gray matter volume and cortical thickness during the early postpartum which relate to positive parenting perceptions (see Kim, 2021). Findings to date suggest that the maternal brain demonstrates both decreases and increases in size which may be temporally and/or regionally specific.

The degree of neural plasticity observed in both animals and humans suggests that stress response neurocircuitry may be altered in pregnancy and lactation, which, coupled with the altered basal activity and reactivity of the HPA axis during this time, could have lasting impacts on the stress response.

6.3 Potential for recalibration

Plasticity in the stress response system and brain are conditions necessary for recalibration. However, what evidence exists that changes in the stress system during pregnancy and lactation extend beyond this period? Cortisol levels appear to return to baseline after the first few months postpartum, and dampened responsivity of the axis is presumed to conclude after weaning. In this sense, alterations in the stress system over pregnancy and lactation can be considered transient. However, recalibration does not mean that the system remains in its pregnant/lactating state for an extended time.

Rather, it would mean that if environmental conditions have changed significantly since an earlier period of recalibration (i.e., puberty), the stress response may recalibrate to match conditions encountered during pregnancy and lactation. It is these changes that would persist.

Several lines of evidence must be established to infer recalibration. First, data are needed showing that significant adversity during pregnancy, especially due to something outside the woman's control, produces alterations in the stress system that last months or years following pregnancy. Some existing studies report that exposure to major/traumatic life events (e.g., Obel et al., 2005) or material deprivation (e.g., Bosquet Enlow et al., 2019) during pregnancy is associated with differences in HPA axis activity, but to our knowledge, no studies have tracked women longitudinally after pregnancy. Such findings would be further strengthened if lasting changes were evident among pregnant/lactating women but not among non-pregnant/lactating women exposed to the same stress.

Of course, the strongest support for recalibration would be evidence that if women experience marked shifts in environments from early life to pregnancy, their stress systems adjust accordingly, and these adjustments persist well beyond pregnancy and lactation. Although there has been a surge of interest in the impacts of early life stress on the HPA axis in pregnancy (see Epstein, Houfek, Rice, & Weiss, 2021), existing study designs do not allow for inferences about changes in the stress response as conditions shift. However, some evidence exists suggesting that supportive conditions during pregnancy may result in different stress system activity for women who experienced harsh conditions early in life. Bublitz and colleagues (Bublitz, Parade, & Stroud, 2014) showed that women with more severe child sexual abuse histories do not exhibit increases in the CAR over gestation seen in other women with such histories if currently experiencing a supportive family environment. Similarly, pregnant women reporting childhood adversity showed higher diurnal cortisol output only if they also reported low current social support, with no association among women with higher support (Thomas, Letourneau, Campbell, & Giesbrecht, 2018). Across these and several other studies (e.g., Stephens et al., 2021), results highlight the possibility that supportive conditions during pregnancy may allow for changes in the stress system so that women who experienced harsh conditions in childhood now show profiles comparable to women with more benign histories.

Findings reviewed above suggest that maternal HPA axis activity during pregnancy and the postpartum reflects both earlier and current environmental conditions. However, it is yet unclear whether harsh (and supportive) environments experienced during pregnancy have stronger and more persisting impacts on the stress response relative to those encountered outside of pregnancy and lactation. This also will be important to demonstrate to infer that pregnancy/lactation opens a particularly sensitive window for recalibration. Because it will be challenging to assess this in humans, translational research programs will be needed to fully test whether recalibration occurs.

Rodent research demonstrates that stress interferes with neural plasticity during pregnancy and lactation, with lasting impacts; these effects do not persist with time in non-pregnant rodents (Slattery & Hillerer, 2016), offering indirect evidence in support of recalibration. Rodent research paradigms also can examine the impacts of environmental enrichment during pregnancy on brain plasticity and the stress response; some have speculated that reproductive experience itself is a form of environmental enrichment (see Pawluski, Lambert, & Kinsley, 2016). Optimally, translational models can align mechanistic rodent studies with longitudinal and intervention studies in humans to draw stronger inferences about recalibration in humans (for a related example, see Morrison et al., 2017).

Several outstanding questions will be important to address to understand if and how pregnancy and lactation open a window for stress recalibration, such as when the window opens and closes. Also, it is unclear whether recalibration would occur only in a first pregnancy, or with each pregnancy. If with each pregnancy, the extent of recalibration may lessen, or may involve slightly different mechanisms, as some neural, hormonal, and behavioral changes differ between first and subsequent pregnancies (see Maupin, Roginiel, Rutherford, & Mayes, 2016). These and other questions can inform the careful design of animal to human translational studies in the search for evidence of recalibration.

6.4 Recalibration and behavior

As in puberty, if the neuroendocrine stress system recalibrates during pregnancy and lactation, it is possible this recalibration extends to other stress-related systems, such as the neurocircuitry involved in precautionary and defensive maternal caregiving behaviors (see Hahn-Holbrook, Holbrook, & Haselton, 2011). Another system involved in parenting behavior (Bos, 2017), the oxytocin (OT)-arginine vasopressin system, may also recalibrate during pregnancy and lactation. Ellis and colleagues

(Ellis, Horn, Carter, van IJzendoorn, & Bakermans-Kranenburg, 2021) recently expanded the ACM to this system, demonstrating meta-analytic evidence that individuals reared in harsh environments are more likely to show lower levels of OT. Given the bidirectional relationship between the OT and stress systems (Alley, Diamond, Lipschitz, & Grewen, 2019), if the HPA system recalibrates in pregnancy, this might support recalibration of the OT system as well.

6.5 Summary

Much research is needed to understand if pregnancy and lactation involve recalibration of the stress response. Work is needed showing that a significant shift in environmental conditions from earlier calibration periods to pregnancy results in adjustments to the stress response that persist well beyond pregnancy and lactation. While not considered here, it is possible that non-pregnant partners or adoptive caregivers experience a degree of stress recalibration over the transition to parenthood; research is emerging regarding paternal brain changes (e.g., Paternina-Die et al., 2020). Periods of stress recalibration may also exist during later adrenal and reproductive hormonal transitions in the lifespan, such as menopause (see Gordon, Eisenlohr-Moul, Rubinow, Schrubbe, & Girdler, 2016). We hope this discussion motivates research into the possibility of recalibration in pregnancy and lactation, which, if established, could have important implications for women's lifespan health.

7. Conclusions and future directions

Stress-responsive physiological systems, most notably the HPA and SAM systems, do appear to be sensitive to the harshness of living conditions early in life. There is increasing evidence in humans that the ex-utero environment affects the developing fetus, which can result in reduced fetal growth and larger HPA axis responses for years after birth. There is also growing evidence that the HPA axis, and perhaps the SAM system, continue to be highly open to environmental input in the first few years after birth. When conditions change markedly, as is the case for children adopted into families from orphanages or orphanage-like institutions, those adopted post-infancy (i.e., approximately 2 years) continue to show the blunted HPA and increased SAM activity observed in children still living in institutional care. Evidence that the stress system is calibrated during fetal and early postnatal life is consistent with a large body of work in animal models that has also identified mechanisms through which calibration is achieved. Nonetheless, there are a number of issues related to these early periods that need resolution, not the least of which is why we tend to see hyper-reactivity with harsh prenatal conditions and hypo-reactivity with harsh postnatal conditions. Is this due to the different nature of the inputs, the different maturational stage of the stress system, or some interaction of both?

Research on recalibration post-infancy is just beginning. The fact that previously-institutionalized children continue to exhibit blunted HPA axis activity throughout the preschool and into the early school years after being adopted in late infancy does not support the early childhood years as being a period of open to recalibration. However, it is not clear that this is the case for the SAM system, as one study of previously institutionalized children showed differences in PEP soon after adoption that in the same children was no longer seen around a year later. Thus, the SAM system may be more open to recalibration across developmental periods than the HPA system. Much more research is needed, however, before this can be concluded.

There is emerging evidence of HPA axis recalibration in during puberty, beginning it seems at the earliest period of adrenarche, but taking many pubertal stages to complete. This has been shown both among previously-institutionalized children in the United States and left-behind children in China. It has also been noted in some studies using the cortisol awakening response. More evidence, however, is needed before concluding that puberty is a period of stress system recalibration. Most importantly, to date, there is no evidence that adults who experience HPA axis recalibration during puberty continue to show stress responses comparable to those of individuals who have lived in more benign, supportive conditions throughout life. These would be very important data, as it is conceivable that shifting from blunted to responsive profiles continues, and individuals who started their lives under harsh, depriving conditions will end up with hyper-responsive systems once the system has fully matured. Furthermore, although the possibility of recalibration is supported by animal model work, there is a disconnect between what is observed in rats and mice during the peripubertal period and what is observed in humans during adrenarche and gonadarche with regards to the HPA axis. It is possible that the increased reactivity of the axis during gonadarche, though, could reflect heightened sensitivity to the type of stress (social evaluation) used most often in human studies. Thus, work using other types of stressors is need before we can confirm that the human pattern of stress system development is different.

While the female neuroendocrine stress system undergoes dynamic changes during pregnancy, we do not yet have evidence that this is a period of stress system recalibration. There is an increased focus on how stressful experiences during pregnancy and during a woman's own childhood affect maternal biology and behavior, namely as pertaining to fetal and infant development. However, there is also growing consideration of how pregnancy and motherhood impact women's life course trajectories. The possibility for pregnancy as a period of stress system recalibration aligns well with this perspective.

Finally, the boundaries of stress recalibration are fuzzy. The pubertal stress recalibration hypothesis is firmly focused on stress-responsive physiological systems. But the possibility that puberty may open a window for recalibration of limbic and behavioral threat-response systems and thus reset risks for psychopathology is sparking a good deal of interest. If recalibration centers on stress-responsive hormonal systems, then changes in other systems would be expected to reflect the impacts of these hormones on brain circuits, especially those that are themselves more plastic during puberty. Evidence, however, that previously institutionalized youth who exhibited a recalibration of the HPA axis towards the normatively responsive pattern shown by comparison youth also showed an increase in internalizing symptoms, suggests that recalibration may not lead to uniformly positive effects on psychological and behavioral functioning. Understanding associations between calibration/recalibration of stress responsive systems, brain circuitry, and behavior is a critical area for future study.

References

Alexander, N., Rosenlöcher, F., Stalder, T., Linke, J., Distler, W., Morgner, J., et al. (2012). Impact of antenatal synthetic glucocorticoid exposure on endocrine stress reactivity in term-born children. *The Journal of Clinical Endocrinology & Metabolism*, *97*(10), 3538–3544. https://doi.org/10.1210/jc.2012-1970.

Alley, J., Diamond, L. M., Lipschitz, D. L., & Grewen, K. (2019). Associations between oxytocin and cortisol reactivity and recovery in response to psychological stress and sexual arousal. *Psychoneuroendocrinology*, *106*, 47–56. https://doi.org/10.1016/j.psyneuen.2019.03.031.

Almanza-Sepulveda, M. L., Fleming, A. S., & Jonas, W. (2020). Mothering revisited: A role for cortisol? *Hormones and Behavior*, *121*, 104679. https://doi.org/10.1016/j.yhbeh.2020.104679.

Avishai-Eliner, S., Brunson, K. L., Sandman, C. A., & Baram, T. Z. (2002). Stressed-out, or in (utero)? *Trends in Neurosciences*, *25*(10), 518–524. https://doi.org/10.1016/S0166-2236(02)02241-5.

Bateson, P., Gluckman, P., & Hanson, M. (2014). The biology of developmental plasticity and the predictive adaptive response hypothesis. *The Journal of Physiology*, *592*(11), 2357–2368. https://doi.org/10.1113/jphysiol.2014.271460.

Batista, G., & Hensch, T. K. (2019). Critical period regulation by thyroid hormones: Potential mechanisms and sex-specific aspects. *Frontiers in Molecular Neuroscience*, *12*. https://www.frontiersin.org/article/10.3389/fnmol.2019.00077.

Bingham, B. C., Rani, C. S. S., Frazer, A., Strong, R., & Morilak, D. A. (2013). Exogenous prenatal corticosterone exposure mimics the effects of prenatal stress on adult brain stress response systems and fear extinction behavior. *Psychoneuroendocrinology*, *38*(11), 2746–2757. https://doi.org/10.1016/j.psyneuen.2013.07.003.

Bos, P. A. (2017). The endocrinology of human caregiving and its intergenerational transmission. *Development and Psychopathology*, *29*(3), 971–999. https://doi.org/10.1017/S0954579416000973.

Bosquet Enlow, M., Sideridis, G., Chiu, Y.-H. M., Nentin, F., Howell, E. A., Grand, B. A. L., et al. (2019). Associations among maternal socioeconomic status in childhood and pregnancy and hair cortisol in pregnancy. *Psychoneuroendocrinology*, *99*, 216–224. https://doi.org/10.1016/j.psyneuen.2018.09.017.

Bublitz, M. H., Parade, S., & Stroud, L. R. (2014). The effects of childhood sexual abuse on cortisol trajectories in pregnancy are moderated by current family functioning. *Biological Psychology*, *103*, 152–157. https://doi.org/10.1016/j.biopsycho.2014.08.014.

Buss, C., Davis, E. P., Shahbaba, B., Pruessner, J. C., Head, K., & Sandman, C. A. (2012). Maternal cortisol over the course of pregnancy and subsequent child amygdala and hippocampus volumes and affective problems. *Proceedings of the National Academy of Sciences*, *109*(20), E1312. https://doi.org/10.1073/pnas.1201295109.

Cao-Lei, L., de Rooij, S. R., King, S., Matthews, S. G., Metz, G. A. S., Roseboom, T. J., et al. (2020). Prenatal stress and epigenetics. *Neuroscience & Biobehavioral Reviews*, *117*, 198–210. https://doi.org/10.1016/j.neubiorev.2017.05.016.

Carmona, S., Martínez-García, M., Paternina-Die, M., Barba-Müller, E., Wierenga, L. M., Alemán-Gómez, Y., et al. (2019). Pregnancy and adolescence entail similar neuroanatomical adaptations: A comparative analysis of cerebral morphometric changes. *Human Brain Mapping*, *40*(7), 2143–2152. https://doi.org/10.1002/hbm.24513.

Colich, N. L., Sheridan, M. A., Humphreys, K. L., Wade, M., Tibu, F., Nelson, C. A., et al. (2021). Heightened sensitivity to the caregiving environment during adolescence: Implications for recovery following early-life adversity. *Journal of Child Psychology and Psychiatry*, *62*(8). https://doi.org/10.1111/jcpp.13347.

Cox, E. Q., Stuebe, A., Pearson, B., Grewen, K., Rubinow, D., & Meltzer-Brody, S. (2015). Oxytocin and HPA stress axis reactivity in postpartum women. *Psychoneuroendocrinology*, *55*, 164–172. https://doi.org/10.1016/j.psyneuen.2015.02.009.

Creutzberg, K. C., Sanson, A., Viola, T. W., Marchisella, F., Begni, V., Grassi-Oliveira, R., et al. (2021). Long-lasting effects of prenatal stress on HPA axis and inflammation: A systematic review and multilevel meta-analysis in rodent studies. *Neuroscience & Biobehavioral Reviews*, *127*, 270–283. https://doi.org/10.1016/j.neubiorev.2021.04.032.

Curran, M. M., Sandman, C. A., Davis, E. P., Glynn, L. M., & Baram, T. Z. (2017). Abnormal dendritic maturation of developing cortical neurons exposed to corticotropin releasing hormone (CRH): Insights into effects of prenatal adversity? *PLoS One*, *12*(6), e0180311. https://doi.org/10.1371/journal.pone.0180311.

Dahl, R. E., & Gunnar, M. R. (2009). Heightened stress responsiveness and emotional reactivity during pubertal maturation: Implications for psychopathology. *Development and Psychopathology*, *21*(1), 1–6. https://doi.org/10.1017/S0954579409000017.

Davis, E. P., Glynn, L. M., Waffarn, F., & Sandman, C. A. (2011). Prenatal maternal stress programs infant stress regulation. *Journal of Child Psychology and Psychiatry*, *52*(2), 119–129. https://doi.org/10.1111/j.1469-7610.2010.02314.x.

Davis, E. P., Waffarn, F., & Sandman, C. A. (2011). Prenatal treatment with glucocorticoids sensitizes the HPA axis response to stress among full-term infants. *Developmental Psychobiology*, *53*(2), 175–183. https://doi.org/10.1002/dev.20510.

de Weerth, C., & Buitelaar, J. K. (2005). Cortisol awakening response in pregnant women. *Psychoneuroendocrinology*, *30*(9), 902–907. https://doi.org/10.1016/j.psyneuen.2005.05.003.

De Wolff, M., & van IJzendoorn, M.H. (1997). Sensitivity and attachment: A meta-analysis on parental antecedents of infant attachment. *Child Development*, *68*(4), 571–591. https://doi.org/10.2307/1132107.

Del Giudice, M., Ellis, B. J., & Shirtcliff, E. A. (2011). The adaptive calibration model of stress responsivity. *Neuroscience & Biobehavioral Reviews*, *35*(7), 1562–1592. https://doi.org/10.1016/j.neubiorev.2010.11.007.

DePasquale, C. E., Donzella, B., & Gunnar, M. R. (2019). Pubertal recalibration of cortisol reactivity following early life stress: A cross-sectional analysis. *Journal of Child Psychology and Psychiatry*, *60*(5), 566–575. https://doi.org/10.1111/jcpp.12992.

Diaz, R., Brown, R. W., & Seckl, J. R. (1998). Distinct ontogeny of glucocorticoid and mineralocorticoid receptor and 11β-hydroxysteroid dehydrogenase types I and II mRNAs in the fetal rat brain suggest a complex control of glucocorticoid actions. *The Journal of Neuroscience*, *18*(7), 2570. https://doi.org/10.1523/JNEUROSCI.18-07-02570.1998.

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, *130*(3), 355–391. https://doi.org/10.1037/0033-2909.130.3.355.

Doom, J. R., Peckins, M. K., Hein, T. C., Dotterer, H. L., Mitchell, C., Lopez-Duran, N. L., et al. (2020). Differential associations of parental harshness and parental disengagement with overall cortisol output at 15 years: Implications for adolescent mental health. *Development and Psychopathology*, *1–18*. https://doi.org/10.1017/S0954579420000954.

Dunn, E. C., Soare, T. W., Zhu, Y., Simpkin, A. J., Suderman, M. J., Klengel, T., et al. (2019). Sensitive periods for the effect of childhood adversity on DNA methylation: Results from a prospective, longitudinal study. *Biological Psychiatry*, *85*(10), 838–849. https://doi.org/10.1016/j.biopsych.2018.12.023.

Eiland, L., & Romeo, R. D. (2013). Stress and the developing adolescent brain. *Neuroscience*, *249*, 162–171. https://doi.org/10.1016/j.neuroscience.2012.10.048.

Ellis, B. J., Horn, A. J., Carter, C. S., van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2021). Developmental programming of oxytocin through variation in early-life stress: Four meta-analyses and a theoretical reinterpretation. *Clinical Psychology Review*, *86*, 101985. https://doi.org/10.1016/j.cpr.2021.101985.

Engel, M. L., & Gunnar, M. R. (2020). The development of stress reactivity and regulation during human development. In A. Clow, & N. Smyth (Eds.), *Vol. 150. Stress and brain health: Across the life course* (pp. 41–76). Academic Press. https://doi.org/10.1016/bs.irn.2019.11.003.

Epstein, C. M., Houfek, J. F., Rice, M. J., & Weiss, S. J. (2021). Integrative review of early life adversity and cortisol regulation in pregnancy. *Journal of Obstetric, Gynecologic & Neonatal Nursing*, *50*(3), 242–255. https://doi.org/10.1016/j.jogn.2020.12.006.

Gabard-Durnam, L. J., & McLaughlin, K. A. (2019). Do sensitive periods exist for exposure to adversity? *Biological Psychiatry*, *85*(10), 789–791. https://doi.org/10.1016/j.biopsych.2019.03.975.

Gatewood, J. D., Morgan, M. D., Eaton, M., McNamara, I. M., Stevens, L. F., Macbeth, A. H., et al. (2005). Motherhood mitigates aging-related decrements in learning and memory and positively affects brain aging in the rat. *Brain Research Bulletin*, *66*(2), 91–98. https://doi.org/10.1016/j.brainresbull.2005.03.016.

George, S. A., Stout, S. A., Tan, M., Knox, D., & Liberzon, I. (2013). Early handling attenuates enhancement of glucocorticoid receptors in the prefrontal cortex in an animal model of post-traumatic stress disorder. *Biology of Mood & Anxiety Disorders*, *3*(1), 22. https://doi.org/10.1186/2045-5380-3-22.

Gitau, R., Fisk, N. M., Teixeira, J. M. A., Cameron, A., & Glover, V. (2001). Fetal hypothalamic-pituitary-adrenal stress responses to invasive procedures are independent of maternal responses. *The Journal of Clinical Endocrinology & Metabolism*, *86*(1), 104–109. https://doi.org/10.1210/jcem.86.1.7090.

Glynn, L. M., Dunkel Schetter, C., Hobel, C. J., & Sandman, C. A. (2008). Pattern of perceived stress and anxiety in pregnancy predicts preterm birth. *Health Psychology*, *27*(1), 43–51. https://doi.org/10.1037/0278-6133.27.1.43.

Glynn, L. M., Howland, M. A., & Fox, M. (2018). Maternal programming: Application of a developmental psychopathology perspective. *Development and Psychopathology*, *30*(3), 905–919. https://doi.org/10.1017/S0954579418000524.

Godoy, L. D., Rossignoli, M. T., Delfino-Pereira, P., Garcia-Cairasco, N., & de Lima Umeoka, E. H. (2018). A comprehensive overview on stress neurobiology: Basic concepts and clinical implications. *Frontiers in Behavioral Neuroscience*, *12*. https://www.frontiersin.org/article/10.3389/fnbeh.2018.00127.

Golm, D., Maughan, B., Barker, E. D., Hill, J., Kennedy, M., Knights, N., et al. (2020). Why does early childhood deprivation increase the risk for depression and anxiety in adulthood? A developmental cascade model. *Journal of Child Psychology and Psychiatry*, *61*(9), 1043–1053. https://doi.org/10.1111/jcpp.13205.

Gordon, J. L., Eisenlohr-Moul, T. A., Rubinow, D. R., Schrubbe, L., & Girdler, S. S. (2016). Naturally occurring changes in estradiol concentrations in the menopause transition predict morning cortisol and negative mood in perimenopausal depression. *Clinical Psychological Science*, *4*(5), 919–935. https://doi.org/10.1177/2167702616647924.

Graham, A. M., Rasmussen, J. M., Entringer, S., Ben Ward, E., Rudolph, M. D., Gilmore, J. H., et al. (2019). Maternal cortisol concentrations during pregnancy and sex-specific associations with neonatal amygdala connectivity and emerging internalizing behaviors. *Biological Psychiatry*, *85*(2), 172–181. https://doi.org/10.1016/j.biopsych.2018.06.023.

Gunnar, M. R. (1992). Reactivity of the hypothalamic-pituitary-adrenocortical system to stressors in normal infants and children. *Pediatrics*, *90*(3), 491–497. https://doi.org/10.1542/peds.90.3.491.

Gunnar, M. R., & Bowen, M. (2021). What was learned from studying the effects of early institutional deprivation. *Pharmacology Biochemistry and Behavior*, *210*, 173272. https://doi.org/10.1016/j.pbb.2021.173272.

Gunnar, M. R., DePasquale, C. E., Reid, B. M., Donzella, B., & Miller, B. S. (2019). Pubertal stress recalibration reverses the effects of early life stress in postinstitutionalized children. *Proceedings of the National Academy of Sciences*, *116*(48), 23984. https://doi.org/10.1073/pnas.1909699116.

Gunnar, M. R., & Reid, B. M. (2019). Early deprivation revisited: Contemporary studies of the impact on young children of institutional care. *Annual Review of Developmental Psychology*, *1*(1), 93–118. https://doi.org/10.1146/annurev-devpsych-121318-085013.

Gunnar, M. R., Wewerka, S., Frenn, K., Long, J. D., & Griggs, C. (2009). Developmental changes in hypothalamus–pituitary–adrenal activity over the transition to adolescence: Normative changes and associations with puberty. *Development and Psychopathology*, *21*(1), 69–85.

Hahn-Holbrook, J., Holbrook, C., & Haselton, M. G. (2011). Parental precaution: Neurobiological means and adaptive ends. *Neuroscience & Biobehavioral Reviews*, *35*(4), 1052–1066. https://doi.org/10.1016/j.neubiorev.2010.09.015.

Handlin, L., Jonas, W., Petersson, M., Ejdebäck, M., Ransjö-Arvidson, A.-B., Nissen, E., et al. (2009). Effects of sucking and skin-to-skin contact on maternal ACTH and cortisol levels during the second day postpartum—Influence of epidural analgesia and oxytocin in the perinatal period. *Breastfeeding Medicine*, *4*(4), 207–220. https://doi.org/10.1089/bfm.2009.0001.

Harteveld, L. M., Nederend, I., ten Harkel, A. D. J., Schutte, N. M., de Rooij, S. R., Vrijkotte, T. G. M., et al. (2021). Maturation of the cardiac autonomic nervous system activity in children and adolescents. *Journal of the American Heart Association*, *10*(4), e017405. https://doi.org/10.1161/JAHA.120.017405.

Hasiec, M., & Misztal, T. (2018). Adaptive modifications of maternal hypothalamic-pituitary-adrenal axis activity during lactation and salsolinol as a new player in this phenomenon. *International Journal of Endocrinology*, *2018*, 3786038. https://doi.org/10.1155/2018/3786038.

Hoekzema, E., Barba-Müller, E., Pozzobon, C., Picado, M., Lucco, F., García-García, D., et al. (2017). Pregnancy leads to long-lasting changes in human brain structure. *Nature Neuroscience*, *20*(2), 287–296. https://doi.org/10.1038/nn.4458.

Howland, M. A., Donzella, B., Miller, B. S., & Gunnar, M. R. (2020). Pubertal recalibration of cortisol-DHEA coupling in previously-institutionalized children. *Hormones and Behavior*, *125*, 104816. https://doi.org/10.1016/j.yhbeh.2020.104816.

Howland, M. A., Sandman, C. A., & Glynn, L. M. (2017). Developmental origins of the human hypothalamic-pituitary-adrenal axis. *Expert Review of Endocrinology & Metabolism*, *12*(5), 321–339. https://doi.org/10.1080/17446651.2017.1356222.

Irwin, J. L., Meyering, A. L., Peterson, G., Glynn, L. M., Sandman, C. A., Hicks, L. M., et al. (2021). Maternal prenatal cortisol programs the infant hypothalamic–pituitary–adrenal axis. *Psychoneuroendocrinology*, *125*, 105106. https://doi.org/10.1016/j.psyneuen.2020.105106.

Katzow, M., Messito, M. J., Mendelsohn, A. L., Scott, M. A., & Gross, R. S. (2019). The protective effect of prenatal social support on infant adiposity in the first 18 months of life. *The Journal of Pediatrics*, *209*, 77–84. https://doi.org/10.1016/j.jpeds.2019.02.017.

Kertes, D. A., Kamin, H. S., Hughes, D. A., Rodney, N. C., Bhatt, S., & Mulligan, C. J. (2016). Prenatal maternal stress predicts methylation of genes regulating the hypothalamic–pituitary–adrenocortical system in mothers and newborns in the Democratic Republic of Congo. *Child Development*, *87*(1), 61–72. https://doi.org/10.1111/cdev.12487.

Kim, P. (2021). How stress can influence brain adaptations to motherhood. *Frontiers in Neuroendocrinology*, *60*, 100875. https://doi.org/10.1016/j.yfrne.2020.100875.

King, L. S., Colich, N. L., LeMoult, J., Humphreys, K. L., Ordaz, S. J., Price, A. N., et al. (2017). The impact of the severity of early life stress on diurnal cortisol: The role of puberty. *Psychoneuroendocrinology*, *77*, 68–74. https://doi.org/10.1016/j.psyneuen.2016.11.024.

King, L. S., Graber, M. G., Colich, N. L., & Gotlib, I. H. (2020). Associations of waking cortisol with DHEA and testosterone across the pubertal transition: Effects of threat-related early life stress. *Psychoneuroendocrinology*, *115*, 104651. https://doi.org/10.1016/j.psyneuen.2020.104651.

Kivlighan, K. T., DiPietro, J. A., Costigan, K. A., & Laudenslager, M. L. (2008). Diurnal rhythm of cortisol during late pregnancy: Associations with maternal psychological well-being and fetal growth. *Psychoneuroendocrinology*, *33*(9), 1225–1235. https://doi.org/10.1016/j.psyneuen.2008.06.008.

Krontira, A. C., Cruceanu, C., & Binder, E. B. (2020). Glucocorticoids as mediators of adverse outcomes of prenatal stress. *Trends in Neurosciences*, *43*(6), 394–405. https://doi.org/10.1016/j.tins.2020.03.008.

Laube, C., van den Bos, W., & Fandakova, Y. (2020). The relationship between pubertal hormones and brain plasticity: Implications for cognitive training in adolescence. *Developmental Cognitive Neuroscience*, *42*, 100753. https://doi.org/10.1016/j.dcn.2020.100753.

Leuner, B., & Sabihi, S. (2016). The birth of new neurons in the maternal brain: Hormonal regulation and functional implications. *Frontiers in Neuroendocrinology*, *41*, 99–113. https://doi.org/10.1016/j.yfrne.2016.02.004.

Love, G., Torrey, N., McNamara, I., Morgan, M., Banks, M., Hester, N. W., et al. (2005). Maternal experience produces long-lasting behavioral modifications in the rat. *Behavioral Neuroscience*, *119*(4), 1084–1096. https://doi.org/10.1037/0735-7044.119.4.1084.

Magiakou, M. A., Mastorakos, G., Rabin, D., Dubbert, B., Gold, P. W., & Chrousos, G. P. (1996). Hypothalamic corticotropin-releasing hormone suppression during the postpartum period: Implications for the increase in psychiatric manifestations at this time. *The Journal of Clinical Endocrinology & Metabolism*, *81*(5), 1912–1917. https://doi.org/10.1210/jcem.81.5.8626857.

Maupin, A. N., Roginiel, A. C., Rutherford, H. J. V., & Mayes, L. C. (2016). A preliminary review of whether prior reproductive experience influences caregiving. *New Directions for Child and Adolescent Development*, *2016*(153), 73–86. https://doi.org/10.1002/cad.20169.

McEwen, B. S. (2017). Allostasis and the epigenetics of brain and body health over the life course: The brain on stress. *JAMA Psychiatry*, *74*(6), 551–552. https://doi.org/10.1001/jamapsychiatry.2017.0270.

McLaughlin, K. A., & Sheridan, M. A. (2016). Beyond cumulative risk: A dimensional approach to childhood adversity. *Current Directions in Psychological Science*, *25*(4), 239–245. https://doi.org/10.1177/0963721416655883.

McLaughlin, K. A., Sheridan, M. A., Tibu, F., Fox, N. A., Zeanah, C. H., & Nelson, C. A. (2015). Causal effects of the early caregiving environment on development of stress response systems in children. *Proceedings of the National Academy of Sciences*, *112*(18), 5637. https://doi.org/10.1073/pnas.1423363112.

Morrison, K. E., Epperson, C. N., Sammel, M. D., Ewing, G., Podcasy, J. S., Hantsoo, L., et al. (2017). Preadolescent adversity programs a disrupted maternal stress reactivity in humans and mice. *Biological Psychiatry*, *81*(8), 693–701. https://doi.org/10.1016/j.biopsych.2016.08.027.

Nazzari, S., Fearon, P., Rice, F., Dottori, N., Ciceri, F., Molteni, M., et al. (2019). Beyond the HPA-axis: Exploring maternal prenatal influences on birth outcomes and stress reactivity. *Psychoneuroendocrinology*, *101*, 253–262. https://doi.org/10.1016/j.psyneuen.2018.11.018.

Nomura, Y., Rompala, G., Pritchett, L., Aushev, V., Chen, J., & Hurd, Y. L. (2021). Natural disaster stress during pregnancy is linked to reprogramming of the placenta transcriptome in relation to anxiety and stress hormones in young offspring. *Molecular Psychiatry*, *26*(11), 6520–6530. https://doi.org/10.1038/s41380-021-01123-z.

Noorlander, C. W., De Graan, P. N. E., Middeldorp, J., Van Beers, J. J. B. C., & Visser, G. H. A. (2006). Ontogeny of hippocampal corticosteroid receptors: Effects of antenatal glucocorticoids in human and mouse. *Journal of Comparative Neurology*, *499*(6), 924–932. https://doi.org/10.1002/cne.21162.

Obel, C., Hedegaard, M., Henriksen, T. B., Secher, N. J., Olsen, J., & Levine, S. (2005). Stress and salivary cortisol during pregnancy. *Psychoneuroendocrinology*, *30*(7), 647–656. https://doi.org/10.1016/j.psyneuen.2004.11.006.

O'Donnell, K. J., & Meaney, M. J. (2020). Epigenetics, development, and psychopathology. *Annual Review of Clinical Psychology*, *16*(1), 327–350. https://doi.org/10.1146/annurev-clinpsy-050718-095530.

Osterholm, E. A., Hostinar, C. E., & Gunnar, M. R. (2012). Alterations in stress responses of the hypothalamic-pituitary-adrenal axis in small for gestational age infants. *Psychoneuroendocrinology*, *37*(10), 1719–1725. https://doi.org/10.1016/j.psyneuen.2012.03.005.

Paternina-Die, M., Martínez-García, M., Pretus, C., Hoekzema, E., Barba-Müller, E., Martín de Blas, D., et al. (2020). The paternal transition entails neuroanatomic adaptations that are associated with the father's brain response to his infant cues. *Cerebral cortex. Communications*, *1*(1), tgaa082. https://doi.org/10.1093/texcom/tgaa082.

Pawluski, J. L., Lambert, K. G., & Kinsley, C. H. (2016). Neuroplasticity in the maternal hippocampus: Relation to cognition and effects of repeated stress. *Hormones and Behavior*, 77, 86–97. https://doi.org/10.1016/j.yhbeh.2015.06.004.

Peper, J. S., & Dahl, R. E. (2013). The teenage brain: Surging hormones—Brain-behavior interactions during puberty. *Current Directions in Psychological Science*, *22*(2), 134–139. https://doi.org/10.1177/0963721412473755.

Perry, N. B., DePasquale, C. E., Donzella, B., & Gunnar, M. R. (2020). Associations between stress reactivity and behavior problems for previously institutionalized youth across puberty. *Development and Psychopathology*, *32*(5), 1854–1863.

Perry, N. B., Donzella, B., & Gunnar, M. R. (2022). Pubertal stress recalibration and later social and emotional adjustment among adolescents: The role of early life stress. *Psychoneuroendocrinology*, *135*, 105578. https://doi.org/10.1016/j.psyneuen.2021.105578.

Petraglia, F., Sutton, S., & Vale, W. (1989). Neurotransmitters and peptides modulate the release of immunoreactive corticotropin-releasing factor from cultured human placental cells. *American Journal of Obstetrics and Gynecology*, *160*(1), 247–251. https://doi.org/10.1016/0002-9378(89)90130-0.

Pluess, M., & Belsky, J. (2011). Prenatal programming of postnatal plasticity? *Development and Psychopathology*, *23*(1), 29–38. https://doi.org/10.1017/S0954579410000623.

Quevedo, K., Johnson, A. E., Loman, M. L., LaFavor, T. L., & Gunnar, M. (2012). The confluence of adverse early experience and puberty on the cortisol awakening response. *International Journal of Behavioral Development*, *36*(1), 19–28. https://doi.org/10.1177/0165025411406860.

Quinn, T., Greaves, R., Badoer, E., & Walker, D. (2018). DHEA in prenatal and postnatal life: Implications for brain and behavior. In G. Litwack (Ed.), *Vitamins and hormones (Vol. 108, pp. 145–174)* Academic Press. https://doi.org/10.1016/bs.vh.2018.03.001.

Reh, R. K., Dias, B. G., Nelson, C. A., Kaufer, D., Werker, J. F., Kolb, B., et al. (2020). Critical period regulation across multiple timescales. *Proceedings of the National Academy of Sciences*, *117*(38), 23242. https://doi.org/10.1073/pnas.1820836117.

Reynolds, R. M. (2013). Glucocorticoid excess and the developmental origins of disease: Two decades of testing the hypothesis – 2012 Curt Richter award winner. *Psychoneuroendocrinology*, *38*(1), 1–11. https://doi.org/10.1016/j.psyneuen.2012.08.012.

Roisman, G. I., Susman, E., Barnett-Walker, K., Booth-LaForce, C., Owen, M. T., Belsky, J., et al. (2009). Early family and child-care antecedents of awakening cortisol levels in adolescence. *Child Development*, *80*(3), 907–920. https://doi.org/10.1111/j.1467-8624.2009.01305.x.

Romeo, R. D. (2017). The impact of stress on the structure of the adolescent brain: Implications for adolescent mental health. *Brain Research*, *1654*, 185–191. https://doi.org/10.1016/j.brainres.2016.03.021.

Romeo, R. D. (2018). The metamorphosis of adolescent hormonal stress reactivity: A focus on animal models. *Frontiers in Neuroendocrinology*, *49*, 43–51. https://doi.org/10.1016/j.yfrne.2017.12.003.

Sandman, C. A. (2018). Prenatal CRH: An integrating signal of fetal distress. *Development and Psychopathology*, *30*(3), 941–952. https://doi.org/10.1017/S0954579418000664.

Seddon, J. A., Rodriguez, V. J., Provencher, Y., Raftery-Helmer, J., Hersh, J., Labelle, P. R., et al. (2020). Meta-analysis of the effectiveness of the Trier social stress test in eliciting physiological stress responses in children and adolescents. *Psychoneuroendocrinology*, *116*, 104582. https://doi.org/10.1016/j.psyneuen.2020.104582.

Simons, S. S. H., Zijlmans, M. A. C., Cillessen, A. H. N., & de Weerth, C. (2019). Maternal prenatal and early postnatal distress and child stress responses at age 6. *Stress*, *22*(6), 654–663. https://doi.org/10.1080/10253890.2019.1608945.

Slattery, D. A., & Hillerer, K. M. (2016). The maternal brain under stress: Consequences for adaptive peripartum plasticity and its potential functional implications. *Frontiers in Neuroendocrinology*, *41*, 114–128. https://doi.org/10.1016/j.yfrne.2016.01.004.

Slattery, D. A., & Neumann, I. D. (2008). No stress please! Mechanisms of stress hyporesponsiveness of the maternal brain. *The Journal of Physiology*, *586*(2), 377–385. https://doi.org/10.1113/jphysiol.2007.145896.

Somerville, L. H., Jones, R. M., Ruberry, E. J., Dyke, J. P., Glover, G., & Casey, B. J. (2013). The medial prefrontal cortex and the emergence of self-conscious emotion in adolescence. *Psychological Science*, *24*(8), 1554–1562. https://doi.org/10.1177/0956797613475633.

Spear, L. P. (2000). The adolescent brain and age-related behavioral manifestations. *Neuroscience & Biobehavioral Reviews*, *24*(4), 417–463. https://doi.org/10.1016/S0149-7634(00)00014-2.

Steine, I. M., LeWinn, K. Z., Lisha, N., Tylavsky, F., Smith, R., Bowman, M., et al. (2020). Maternal exposure to childhood traumatic events, but not multi-domain psychosocial stressors, predict placental corticotrophin releasing hormone across pregnancy. *Social Science & Medicine*, *266*, 113461. https://doi.org/10.1016/j.socscimed.2020.113461.

Stephens, J. E., Kessler, C. L., Buss, C., Miller, G. E., Grobman, W. A., Keenan-Devlin, L., et al. (2021). Early and current life adversity: Past and present influences on maternal diurnal cortisol rhythms during pregnancy. *Developmental Psychobiology*, *63*(2), 305–319. https://doi.org/10.1002/dev.22000.

Stroud, L., Foster, E., Handwerger, K., Papandonatos, G. D., Granger, D., Kivlighan, K. T., & Niaura, R. (2009). Stress response and the adolescent transition: Performance versus peer rejection stress. *Development & Psychopathology*, *21*, 47–68.

Stroud, L. R., Papandonatos, G. D., Williamson, D. E., & Dahl, R. E. (2004). Sex differences in the effects of pubertal development on responses to a corticotropin-releasing hormone challenge: The Pittsburgh psychobiologic studies. *Annals of the New York Academy of Sciences*, *1021*(1), 348–351. https://doi.org/10.1196/annals.1308.043.

Szutorisz, H., & Hurd, Y. L. (2016). Feeding the developing brain: The persistent epigenetic effects of early life malnutrition. *Biological Psychiatry*, *80*(10), 730–732. https://doi.org/10.1016/j.biopsych.2016.08.032.

Tarullo, A. R., Tuladhar, C. T., Kao, K., Drury, E. B., & Meyer, J. (2020). Cortisol and socioeconomic status in early childhood: A multidimensional assessment. *Development and Psychopathology*, *32*(5), 1876–1887. https://doi.org/10.1017/S0954579420001315.

Thayer, Z. M., Wilson, M. A., Kim, A. W., & Jaeggi, A. V. (2018). Impact of prenatal stress on offspring glucocorticoid levels: A phylogenetic meta-analysis across 14 vertebrate species. *Scientific Reports*, *8*(1), 4942. https://doi.org/10.1038/s41598-018-23169-w.

Thomas, J. C., Letourneau, N., Campbell, T. S., & Giesbrecht, G. F. (2018). Social buffering of the maternal and infant HPA axes: Mediation and moderation in the intergenerational transmission of adverse childhood experiences. *Development and Psychopathology*, *30*(3), 921–939. https://doi.org/10.1017/S0954579418000512.

Thomason, M. E. (2020). Development of brain networks in utero: Relevance for common neural disorders. *Biological Psychiatry*, *88*(1), 40–50. https://doi.org/10.1016/j.biopsych.2020.02.007.

van den Bos, E., de Rooij, M., Miers, A. C., Bokhorst, C. L., & Westenberg, P. M. (2014). Adolescents' increasing stress response to social evaluation: Pubertal effects on cortisol and alpha-amylase during public speaking. *Child Development*, *85*(1), 220–236. https://doi.org/10.1111/cdev.12118.

Vermeer, H. J., & van IJzendoorn, M. H. (2006). Children's elevated cortisol levels at daycare: A review and meta-analysis. *Early Childhood Research Quarterly*, *21*(3), 390–401. https://doi.org/10.1016/j.ecresq.2006.07.004.

Vickers, M. H. (2022). Early life nutrition and neuroendocrine programming. *Neuropharmacology*, *205*, 108921. https://doi.org/10.1016/j.neuropharm.2021.108921.

Wade, M., Sheridan, M. A., Zeanah, C. H., Fox, N. A., Nelson, C. A., & McLaughlin, K. A. (2020). Environmental determinants of physiological reactivity to stress: The interacting effects of early life deprivation, caregiving quality, and stressful life events. *Development and Psychopathology*, *32*(5), 1732–1742. https://doi.org/10.1017/S0954579420001327.

Watamura, S. E., Donzella, B., Kertes, D. A., & Gunnar, M. R. (2004). Developmental changes in baseline cortisol activity in early childhood: Relations with napping and effortful control. *Developmental Psychobiology*, *45*(3), 125–133. https://doi.org/10.1002/dev.20026.

Welberg, L. A. M., & Seckl, J. R. (2001). Prenatal stress, glucocorticoids and the programming of the brain. *Journal of Neuroendocrinology*, *13*(2), 113–128. https://doi.org/10.1111/j.1365-2826.2001.00601.x.

Wells, R. G. (2015). The adrenal glands. In *Vol. 1–book, section. Diagnostic imaging of infants and children* McGraw-Hill Education. https://accesspediatrics.mhmedical.com/content.aspx?aid=1148952486.

Xu, Y.-J., Sheng, H., Wu, T.-W., Bao, Q.-Y., Zheng, Y., Zhang, Y.-M., et al. (2018). CRH/CRHR1 mediates prenatal synthetic glucocorticoid programming of depression-like behavior across 2 generations. *The FASEB Journal*, *32*(8), 4258–4269. https://doi.org/10.1096/fj.201700948RR.

Yim, I. S., Quas, J. A., Cahill, L., & Hayakawa, C. M. (2010). Children's and adults' salivary cortisol responses to an identical psychosocial laboratory stressor. *Psychoneuroendocrinology*, *35*(2), 241–248. https://doi.org/10.1016/j.psyneuen.2009.06.014.

Yim, I. S., Quas, J. A., Rush, E. B., Granger, D. A., & Skoluda, N. (2015). Experimental manipulation of the Trier social stress test-modified (TSST-M) to vary arousal across development. *Psychoneuroendocrinology*, *57*, 61–71. https://doi.org/10.1016/j.psyneuen.2015.03.021.

Yong Ping, E., Laplante, D. P., Elgbeili, G., Hillerer, K. M., Brunet, A., O'Hara, M. W., et al. (2015). Prenatal maternal stress predicts stress reactivity at 2½ years of age: The Iowa flood study. *Psychoneuroendocrinology*, *56*, 62–78. https://doi.org/10.1016/j.psyneuen.2015.02.015.

Yong Ping, E., Laplante, D. P., Elgbeili, G., Jones, S. L., Brunet, A., & King, S. (2020). Disaster-related prenatal maternal stress predicts HPA reactivity and psychopathology in adolescent offspring: Project ice storm. *Psychoneuroendocrinology*, *117*, 104697. https://doi.org/10.1016/j.psyneuen.2020.104697.

Young, E. S., Doom, J. R., Farrell, A. K., Carlson, E. A., Englund, M. M., Miller, G. E., et al. (2021). Life stress and cortisol reactivity: An exploratory analysis of the effects of stress exposure across life on HPA-axis functioning. *Development and Psychopathology*, *33*(1), 301–312. https://doi.org/10.1017/S0954579419001779.

Zhang, D., Fang, J., Zhang, L., Yuan, J., Wan, Y., Su, P., et al. (2021). Pubertal recalibration of cortisol reactivity following early life parent-child separation. *Journal of Affective Disorders*, *278*, 320–326. https://doi.org/10.1016/j.jad.2020.09.030.

Zijlmans, M. A. C., Riksen-Walraven, J. M., & de Weerth, C. (2015). Associations between maternal prenatal cortisol concentrations and child outcomes: A systematic review. *Neuroscience & Biobehavioral Reviews*, *53*, 1–24. https://doi.org/10.1016/j.neubiorev.2015.02.015.

This page intentionally left blank

CHAPTER THREE

Parental sexual orientation, parental gender identity, and the development of children

Charlotte J. Patterson*
University of Virginia, Charlottesville, VA, United States
*Corresponding author: e-mail address: cjp@virginia.edu

Contents

Abstract

In recent years, many lesbian, gay, bisexual, transgender, and queer (LGBTQ+) adults have become parents. LGBTQ+ parenthood does, however, remain a controversial topic across the United States and around the world. Several questions have been raised. For instance, to what extent do LGBTQ+ adults make capable parents? Do children who have LGBTQ+ parents grow up in healthy ways? What factors contribute to positive family functioning in families with LGBTQ+ parents? A growing body of social science research has addressed these questions, and the findings suggest both that LGBTQ+ adults are successful in their roles as parents and that their children develop in positive ways. Overall, the findings to date suggest that parental sexual orientation and gender identity do not in themselves determine success in parenting or child development; indeed, sexual and gender minority parents and their children have shown remarkable resilience, even in the face of many challenges. Contextual issues, as well as implications of research findings for law and policy around the world are discussed.

Many lesbian, gay, bisexual, transgender and queer (LGBTQ+) adults want to become parents, and many have succeeded in doing so (Patterson & Riskind, 2010; Reczek, 2020; Riskind & Patterson, 2010; Riskind

Advances in Child Development and Behavior, Volume 63
ISSN 0065-2407
https://doi.org/10.1016/bs.acdb.2022.03.002

& Tornello, 2017; Simon, Tornello, Farr, & Bos, 2018; Stotzer, Herman, & Hasenbush, 2014). Indeed, many lesbian and gay adults have had children who are biologically linked to them, and others have adopted children or become foster parents (Farr, Vazquez, & Patterson, 2020; Gates, 2013; Patterson & Tornello, 2011; Perrin, Hurley, Mattern, Flavin, & Pinderhughes, 2019) According to data from national surveys in the U.S., more than 100,000 LGBTQ+ couples are currently rearing children under 18 years of age (Goldberg & Conron, 2018), and this is almost certainly a significant undercount. There is, however, continued controversy about parenting by LGBTQ+ adults (Patterson, Farr, & Goldberg, 2021). Over the last three decades, a growing body of research on children by LGBTQ+ parents has emerged to address questions at the center of public discussions.

In the context of controversies about LGBTQ+ parenting, this chapter provides an overview of recent research in this area. Much of the work addressing non-heterosexual parent families has focused on lesbian and/or gay parents—almost to the exclusion of other sexual and gender minority identities, such as bisexual and transgender parents (Goldberg, Gartrell, & Gates, 2014; Lamb, 2012; Moore & Stambolis-Ruhstorfer, 2013; Patterson, 2017). For this reason, discussion below will focus mainly on lesbian and gay parents, and on their children. Wherever possible, however, research on parents who identify as bisexual, transgender, or other related identities and their children will also be discussed.

This chapter views the research on LGBTQ+ parents and their children through interdisciplinary and intersectional lenses. Studies of LGBTQ+ parenting have emerged primarily from psychology, but research from social work, family science, demography, sociology, public policy, law, and other disciplines has also been important. For this reason, the body of work emerging from these disparate fields is considered, as well as that from psychology. In addition, most research on LGBTQ+ parenting has focused primarily on parental sexual and gender identities (Fish & Russell, 2018). In this review, where possible, other intersecting identities, such as race, class, and legal or policy contexts are considered. It should be acknowledged, however, that research on people of color and/or on those living in poverty has been relatively scarce (Reczek, 2020). Within these contexts, then, the available research on LGBTQ+ parents and their children is presented below. Throughout the chapter, both similarities and differences among LGBTQ+ and cisgender heterosexual parent families are considered. Findings that are specific to processes among LGBTQ+ parent families, such as talking with children about having LGBTQ+ parents are also discussed.

Given that research on these topics began in the English-speaking world and developed for many years in that context—for example, in the U.S., and in the U.K.—the discussion will focus mainly on research findings from these parts of the world. The importance of the legal and cultural contexts in which families live must nevertheless be recognized. There has been a recent explosion of research on LGBTQ+ parenting from around the world, in varied legal and cultural contexts, and a brief summary of the results of this work will also be presented. This will be followed by conclusions and discussion about directions for future research.

1. Early controversies and research about LGBTQ+ parenting

For many years, parenting by LGBTQ+ adults has been a controversial issue, both in the U.S. and around the world (Golombok, 2015; Lamb, 2012; Patterson, 1992, 2006, 2009; Patterson et al., 2021). In this regard, research and law have been intertwined in many ways. For instance, the ways in which lesbian and gay parents have historically been discriminated against in U.S. courts of law have provided one important impetus for research. At the same time, research has played a significant role in support of legal change over time. Because they form a significant part of the context of LGBTQ+ parenting, some of the major legal policies in the U.S. relevant to sexual and gender minority parents are identified briefly, together with a few of the ways in which research may have informed the courts' recent consideration of them.

In context of widely held assumptions about the inherent superiority of traditional family structures, many observers once expected that children of lesbian or gay parents would suffer from a number of problems (Baumrind, 1995; Falk, 1989; Hitchens & Kirkpatrick, 1985). For instance, lesbian and gay parents were seen as unlikely to succeed in parental roles, and as inappropriate role models for children. Moreover, it was expected that children would be teased, harassed, and rejected by their peers, and that this would cause psychological problems. It was also argued that the children of lesbian and gay parents might show atypical gender development. Some observers wondered if children might even grow up to be lesbian or gay themselves, an outcome that—from a heteronormative perspective—was generally considered to be negative (Falk, 1989).

Unproven notions like these were behind many of the discriminatory laws that, for many years, did real damage to many lesbian and gay parents, and to their children. In many parts of the U.S., parents were denied custody of their children, and prospective parents were denied the opportunity to become foster parents or to adopt children; in some jurisdictions, second parents in same-sex couples were also denied the possibility of adopting children they had reared from birth (Ball, 2012; Patterson, 1992, 1995b; Rivers, 2013). Since most of the negative ideas were in essence empirical hypotheses, however, they also gave rise to a considerable body of research that was designed to evaluate them (Goldberg & Allen, 2020; Golombok, 2015; Lamb, 2012; Patterson, 1992, 2000, 2006, 2017).

In an important example of the early research, Golombok, Spencer, and Rutter (1983) studied children of divorced lesbian mothers, comparing them to same-aged children of divorced heterosexual mothers. Their study revealed no significant differences in children's gender development, behavior problems, emotional difficulties, peer relations, or social development as a function of parental sexual orientation; moreover, children's adjustment was not associated with parental sexual orientation (Golombok et al., 1983). Many other researchers published similar results in studies of divorced lesbian mothers, divorced gay fathers, and women who had children after coming out as lesbian (Patterson, 1992, 2000, 2017). As research findings of this kind began to reach the general public as well as the courts, and as public attitudes became more accepting (Patterson, Sepulveda, & White, 2020), judges began to render different kinds of decisions. Over time, the legal situation of lesbian and gay parents in the U.S. gradually grew more favorable (Ball, 2012),

Significant decisions relevant to sexual orientation and parenting were made in courts across the country, and at least some of these were informed by research findings. One such decision from the U.S. Supreme Court was Lawrence v. Texas (2003), which overturned so-called sodomy laws across the country. Sodomy laws, which forbade oral or anal sex, and in some states made consensual same-sex sexual behavior into a felony offense, had often been used to characterize lesbian and gay parents as criminals, and thus unfit to be parents (Ball, 2012). After the *Lawrence* decision, it became more difficult to deny custody of their children to LGBTQ+ parents. Another landmark decision was Obergefell v. Hodges (2015), which legalized same-sex marriages across the country. Among other effects, the *Obergefell* decision opened up expanded legal recognition of parentage, such as step-parent adoptions, in families headed by same-sex parents. In both of these cases, questions about parenting and children were prominent, and many observers have remarked that the findings of social science research seemed to play a role in informing decisions (Ball, 2012; Falk, 1989; Rivers, 2013).

In the U.S. today, the marriages of same-sex couples are recognized by law in all 50 states. A child born to two married same-sex parents is presumed to be the legal child of both parents, regardless of the presence or absence of genetic links. In all 50 states, married same sex couples are allowed to petition for step-parent or joint adoption, so that both petitioners can be recognized as a child's legal parents (Shapiro, 2020). Where once there was widespread discrimination, such that lesbian or gay parents were often at risk of losing custody of their children in the courts, today the law protects the marriages and children of same-sex parents.

One way in which discrimination against lesbian and gay parents is still a reality, however, is in acceptance of so-called "religious exemptions." Even in jurisdictions that have non-discrimination laws in place, religiously-based child welfare agencies have sometimes argued that religious beliefs justify negative treatment of sexual minority parents or prospective parents. Several states (e.g., Kansas, Georgia, Oklahoma) have passed legislation that allows state-licensed child welfare agencies to refuse services to prospective sexual minority foster and adoptive parents based on their religious beliefs (Movement Advancement Project, 2018). In Fulton v. Philadelphia (2021), despite presentation of empirical evidence about the damage that it was likely to cause, the U.S. Supreme Court allowed a religious exemption to stand. The decision was, however, narrowly framed and may not have wide applicability, but this remains to be seen (for further discussion of this case, see Patterson & Farr, 2022).

Thus, for many years, research and law about lesbian and gay parents and their children have been closely connected, and the changing legal situation has formed an important part of the context of LGBTQ+ parenting in the U.S. (Patterson, 1995a, 1995b, 2009). Even as research has contributed to legal change (Gilfoyle & Dvoskin, 2017), controversies surrounding the parenting of children by sexual minority adults have prompted research about the capabilities of lesbian and gay adults as parents, about various outcomes for children reared by lesbian and gay parents, and about overall family interactions and processes in such families. This research is discussed below.

2. Research on LGBTQ+-parent families

In this section, findings of research on how LGBTQ+ adults become parents, their strengths and challenges, and outcomes for children and parents are presented. In some respects, LGBTQ+ parents have experiences that are very similar to those of other parents (e.g., learning to respond to toddlers' tantrums), but they also face some issues that are specific to their circumstances (e.g., questions about which member of a lesbian couple should become pregnant). In what follows, both of these types of issues are considered.

2.1 Pathways to parenthood

LGBTQ+ adults become parents in many different ways (Dickey, Ducheny, & Ehrbar, 2016; Patterson & Riskind, 2010; Perrin et al., 2019; Tornello & Bos, 2017). Some have children in the context of a heterosexual marriage that later dissolves when one or both partners comes out as LGBTQ+; after divorce, a non-heterosexual or transgender parent may then take on either custodial or non-custodial roles for children. Others pursue pathways to parenthood that include adoption, donor insemination, and surrogacy. Each of these pathways is considered briefly below.

Before discussing pathways to parenthood among LGBTQ+ people, however, one should pause to consider some findings about how choices among various pathways may be changing over time. The best evidence on this topic comes from studies of gay fathers. In a survey of more than 700 gay fathers from across the U.S., Tornello and Patterson (2015) reported that most men over 50 years of age had become fathers in context of a heterosexual marriage, whereas most men under 50 had become fathers after coming out, via adoption, surrogacy, or foster care. Similarly, Perrin and her colleagues (2019) also surveyed a large sample of gay fathers in the U.S., and found that, while most of their children who had been born before 1996 had been conceived in the context of heterosexual relationships, most of those born in 2004 or later had joined the gay father's family via adoption, surrogacy, or foster care. These findings illustrate the kinds of changes over time that have been made among gay men as they pursue fatherhood. Trends among sexual minority women are believed to show similar if not larger change in the direction of the pursuit of motherhood after coming out, but systematic studies have not been reported. Overall, the available information suggests that gay men and probably also sexual minority women are increasingly choosing to become parents after coming out. For this reason, adoption, foster care, donor insemination, and surrogacy are being employed more and more frequently among sexual minority adults who seek to become parents.

LGBTQ+ adults may adopt children for reasons that are both similar to, and distinct from, those of cisgender heterosexual adults (Goldberg, 2012; Mallon, 2011; Tornello & Bos, 2017). In Farr and Patterson's (2009) study of adoptive families in the U.S. (about half of whom were lesbian or gay and about half heterosexual), virtually all couples gave "wanted to have children" as a reason for pursuing adoption, regardless of parental sexual orientation. The majority of heterosexual couples reported "challenges with infertility" as another motivation for adopting children, but fewer than half of same-sex couples reported this. On the other hand, many more same-sex than other-sex couples reported that they "did not have a strong desire for biological children." Similarly, in other studies with U.S. samples of lesbian and heterosexual adoptive couples, lesbian women have been less likely than heterosexual women to report a commitment to biological parenthood, attempts to conceive, or pursuit of fertility treatments (Goldberg, Downing, & Richardson, 2009; Goldberg & Smith, 2008).

Many lesbian women and gay men have pursued adoption as a pathway to parenthood (Farr, Vazquez, & Patterson, 2020). Heterosexual adoptive parents have often described adoption as a "second choice" pathway to parenthood, chosen only after struggles with infertility convinced them that biological parenthood was not a realistic option (e.g., Mallon, 2011). In a sample of lesbian, gay, and heterosexual adoptive parent couples in the U.K., Jennings and colleagues reported that both gay and lesbian adoptive parents were less likely than heterosexual adoptive parents to desire, value, or attempt to have a biologically related child (Jennings, Mellish, Tasker, Lamb, & Golombok, 2014). Some transgender adults also describe adoption as a preferred pathway to parenthood (Dickey et al., 2016; Farr & Goldberg, 2018; Tornello & Bos, 2017). Thus, LGBTQ+ adoptive parents are more likely than heterosexual parents to have chosen adoption as their first-choice routes to parenthood (Mallon, 2011).

LGBTQ+ adults may also be more willing than their heterosexual peers to adopt a child from a racial-ethnic background different than their own. Transgender adults may also be especially open to adopting older children or those who are transgender (Goldberg, Tornello, Farr, Smith, & Miranda, 2020). Among preadoptive couples, lesbian couples have been found to be more open than heterosexual couples to transracial adoption (Goldberg, 2009a). Other investigators have also reported that both lesbian and gay adoptive couples are more likely than heterosexual adoptive couples to have completed a transracial adoption (Farr & Patterson, 2009; Lavner, Waterman, & Peplau, 2012; Raleigh, 2012). However, in another study of lesbian, gay, and heterosexual adoptive parents, there were no significant differences in the likelihood of completing a transracial adoption as a function of sexual orientation (Brodzinsky & Goldberg, 2016). Thus, discrepancies in completion rates for transracial adoptions by lesbian, gay, and heterosexual couples warrant further study.

One reason that LGBTQ+ couples in the U.S. may be particularly willing to adopt transracially is that same-sex couples are themselves more likely than heterosexual couples to be interracial, and, in turn, interracial couples are more likely than same-race couples to complete transracial adoptions (Farr & Patterson, 2009; Raleigh, 2012). Indeed, LGBTQ+ parents tend to live in communities with considerable racial diversity within the U.S. (Gates, 2013), which may increase levels of comfort in interracial interactions. Because they are often less committed than heterosexual couples to achieving biological parenthood, LGBTQ+ couples in the U.S. and U.K. may also be more open than heterosexual couples to transracial adoptions (Dickey et al., 2016; Farr & Patterson, 2009; Goldberg et al., 2009; Jennings et al., 2014).

Another way that LGBTQ+ adoptive couples may differ from heterosexual adoptive couples is in terms of gender preferences in adoption. Goldberg (2009b) studied lesbian, gay, and heterosexual couples in the U.S. who were seeking to adopt, and reported that, while heterosexual men were unlikely to express a gender preference, gay men often preferred to adopt boys. Lesbian participants who expressed a preference generally preferred to adopt girls, as did the heterosexual women in the sample. Similar findings have emerged from other research conducted in the U.S. and in Europe regarding preferences for child gender among lesbian, gay, and heterosexual adoptive couples (e.g., Baccara, Collard-Wexler, Felli, & Yariv, 2014; Herrmann-Green & Gehring, 2007).

What might account for gender preferences? In one study, many lesbian and gay adoptive parents explained their preferences for child gender by reference to concerns about gender socialization (Goldberg, 2009b). Some participants reported feeling uncertain about parenting a child whose gender was different than their own. It is possible that lesbian and gay couples, being composed of two parents of the same gender, may have felt uncertain about their ability to parent a child of a different gender. Heterosexual couples, on the other hand, may be less likely to question their ability to parent a child of either gender since one parent of each gender is represented in the parenting couple. Results of research on actual parenting abilities do not sustain such concerns, however, and additional research is needed.

Research has also begun to explore the views of adoptive families in the U.S. about arrangements on the nature and amount of contact between adoptive and birth families, which is often referred to as "openness" in adoption (Farr & Goldberg, 2015). Preliminary research suggests that, as compared to heterosexual adoptive parents, same-sex adoptive parents may be more open to contact with birth relatives (Goldberg, Kinkler, Richardson, & Downing, 2011), and may report more positive relationships with them (Brodzinsky & Goldberg, 2016). As more and more adoptions are characterized by some degree of openness, further research on this topic would be valuable.

Of course, many LGBTQ+ people choose to have children via methods that yield a biological connection between the child and a parent. Sexual minority women may choose to use techniques associated with assisted reproductive technology such as donor insemination, and sexual minority men may choose to pursue surrogacy. Due to legal and financial constraints, however, these pathways are not available to all.

Less is known overall about transgender parents and their children, as compared to lesbian or gay parents, but some information is available. Grant, Mottet, and Tanis (2011), reported on results of the National Transgender Survey, which involved more than 6000 participants, from across the U.S. They found that 38% of transgender adults described themselves as parents. Those who had transitioned later in the life course were more often parents, suggesting that many children had been born in context of heterosexual relationships prior to transition. Those who had been assigned male at birth and transitioned to become women were more likely to describe themselves as parents than those who had been assigned female at birth and transitioned to become men (Grant et al., 2011; Hafford-Letchfield et al., 2019). Additional research on transgender parenting would be valuable (Grant et al., 2011; Tornello & Bos, 2017).

2.2 Challenges and strengths of LGBTQ+ parents

LGBTQ+ adults can face a number of institutional and attitudinal barriers in the process of becoming parents. For instance, not all adoption agencies and/or adoption workers are willing to work with openly LGBTQ+ prospective parents. Brodzinsky (2011) found that, among a large group of public and private adoption agencies throughout the U.S., most (60%) of reporting agencies had accepted applications from lesbian or gay prospective adoptive parents, but only a minority (39%) had actually placed children with lesbian or parents. Jewish, Lutheran, and private non-religious agencies, as well as public agencies or those with a focus on special needs adoption, were most willing to work with lesbian or gay parents. Agencies that were affiliated with conservative religiously affiliated groups (e.g., Baptist, Mormon) were less likely than others to work with LGBTQ+ parents. Over and above outright refusal, discriminatory treatment from adoption workers is a continuing issue among LGBTQ+ prospective parents (Brodzinsky, 2011; Goldberg, Moyer, Kinkler, & Richardson, 2012; Kinkler & Goldberg, 2011; Mallon, 2011; Ross et al., 2008; Ross, Epstein, Anderson, & Eady, 2009; Stotzer et al., 2014). Transgender parents in the U.S. have also reported experiences of discrimination and fear of bias during the adoption process, such as concerns about whether they would be treated fairly if they were open about transgender identities (Farr & Goldberg, 2018; Pyne, 2012; Stotzer et al., 2014).

LGBTQ+ adults seeking to become parents through other pathways may also encounter challenges. Some clinics or health practitioners may not want to work with LGBTQ+ clients in the context of surrogacy or other forms of assisted reproductive technology. Women wanting to become pregnant must obtain sperm, either from known or unknown donors, through a clinic or elsewhere. Men who choose surrogacy also encounter many questions about how to carry out this plan. In recent years, discriminatory practices have become less common, but financial barriers still remain (Berkowitz, 2020; Bos & Gartrell, 2020)

Even though they often need to overcome barriers to parenthood, LGBTQ+ individuals and couples may also offer special strengths in their roles as parents. For example, same sex couples have often been found to divide the labor involved in childcare in a more egalitarian fashion (i.e., more evenly) than do heterosexual couples, and in this way, they provide favorable role models for their children (Bos, van Balen, & van den Boom, 2007; Chan, Brooks, Raboy, & Patterson, 1998; Farr & Patterson, 2013; Patterson, 1995a; Patterson, Sutfin, & Fulcher, 2004). Moreover, among same-sex couples, shared parenting is associated with greater couple relationship adjustment; those who have relatively egalitarian arrangements of childcare are also happier in their relationships (Farr & Patterson, 2013; Patterson et al., 2004). Many researchers have reported similar findings but a few have noted less egalitarian patterns of childcare arrangements among lesbian couples. Some of these findings have come from studies of families in which a woman had children and only later became involved in a relationship with another woman; in these families, the birth mother was likely to report doing more childcare (e.g., Moore, 2011). Overall, the bulk of findings suggest that lesbian and perhaps also gay parenting couples are more likely than their heterosexual counterparts to share childcare labor in a more even fashion than their heterosexual peers (Farr & Patterson, 2013).

With regard to family interaction, lesbian mothers in one study were observed to be more supportive of one another in observations of triadic (i.e., parent/parent/child) interaction among adoptive couples and their children than were heterosexual or gay parents and their children (Farr & Patterson, 2013). Among all family types, more supportive interaction was associated with positive adjustment for children in this sample (Farr & Patterson, 2013). Further research on these and related topics would be helpful.

Regardless of sexual orientation, many individuals report great satisfaction in being a parent. For example, in a study of heterosexual, gay, and lesbian parents of children adopted from foster care in the U.S., Lavner, Waterman, and Peplau (2014) found that parents generally reported satisfaction with their adoption, reported few depressive symptoms, and little parental stress, across three time points (i.e., 2, 12, and 24 months postplacement). Indeed, many LGBTQ+ parents report receiving more support from families of origin than they had expected, and describe themselves as feeling very satisfied with their experience of parenthood (Bergman, Rubio, Green, & Padron, 2010; Brown, Smalling, Groza, & Ryan, 2009; Goldberg & Smith, 2014; Tornello, Kruczkowski, & Patterson, 2015). Thus, not only do LGBTQ+ adults who become parents demonstrate a variety of distinct and unique strengths in these roles, but they also are likely to express satisfaction in their parenting roles.

2.3 The transition to parenthood among LGBTQ+ adults

Regardless of parents' gender or sexual identities, the transition to parenthood brings both joys and challenges. When couples become parents, there is a period of adjustment that can be marked by stress and compromised mental and physical health as well as by happiness and excitement (Cowan & Cowan, 1992; McKay, Ross, & Goldberg, 2010). Consistent with the broader literature on the transition to parenthood, Goldberg, Smith, and Kashy (2010) found that, among lesbian, gay, and heterosexual adoptive couples in the U.S., relationship quality declined across the transition to parenthood; this was true for all types of couples. Women reported the greatest declines in love, and those in relationships with women (i.e., both heterosexual and lesbian partners) reported the greatest ambivalence. In another study based on data from the same sample, Goldberg and

Smith (2009) also found that most parents reported increases in perceived parenting skill across the transition to parenthood.

In a longitudinal study examining factors affecting lesbian and gay adoptive couples across the transition to parenthood, Goldberg and Smith (2008, 2011) found that greater perceived social support and better relationship quality were associated with more favorable mental health for all couples. Sexual minority parents who had higher levels of internalized homophobia and who lived in areas with unfavorable legal climates reported the largest increases in anxiety and depression across the transition to parenthood. In this research, the factors that contributed most to parental well-being during the transition to parenthood were the presence of social support, and other variables related to family processes, not the parents' sexual or gender identity (Goldberg, Kinkler, Moyer, & Weber, 2014; Lavner et al., 2014; Sumontha, Farr, & Patterson, 2016).

With the birth or adoption of a child, one important set of decisions that parents must make involves the choice of children's names. Interesting differences may emerge in this area, as a function of parental sexual orientation. In their study of lesbian, gay, and heterosexual adoptive parents in the U.S., Patterson and Farr (2017) found that heterosexual couples were more likely than lesbian or gay couples to follow patronymic conventions in naming their children. Thus, whereas children of heterosexual parents were most likely to have been given the last names of their fathers, children of lesbian or gay parents were more often given hyphenated last names that had been created by combining the last names of both parents. Thus, same- and other-sex couples in this study took different approaches to naming their children (Patterson & Farr, 2017). Little additional information is available about naming of children by sexual and gender minority parents, and this too is a topic that would benefit from further study (see Pilcher, 2017).

It is worth noting that the transition to parenthood has been studied more carefully among some LGBTQ+ parents than among others. In particular, the special issues of LGBTQ+ couples becoming parents through surrogacy have as yet received less systematic study than others. Like the experiences of adoptive parents in some ways (e.g., neither intended parent has given birth), they are nevertheless very different in other ways (e.g., one parent is often genetically linked with the child, while the other is not). Despite a handful of studies of gay men's transition to parenthood in families formed via surrogacy (e.g., Bergman et al., 2010; Golombok, Blake, Casey, Roman, & Jadva, 2018), transitions to parenthood among same-sex couples using surrogacy are still relatively little studied.

2.4 Parental sexual orientation and child development

In controversies surrounding LGBTQ+ parenting, debate has often centered on children's development (Golombok, 2015; Patterson, 2000, 2017). Questions have been raised about whether LGBTQ+ adults are likely to provide children with adequate parenting, appropriate role models, and effective socialization, particularly in the areas of gender and sexual development. Research on sexual orientation and parenting has been informative here; children of LGBTQ+ parents in general appear to develop in similar ways to those with heterosexual parents (Biblarz & Stacey, 2010; Moore & Stambolis-Ruhstorfer, 2013; Patterson, 2013, 2017). Studies of LGBTQ+ parent families, focusing on children's behavioral adjustment, gender development, and lived experiences related to adoptive and racial-ethnic identity development are reviewed below. Results of research on parenting, couple relationships, parent-child relationships, and on family systems are also discussed. Results of these studies are unusually clear, and they indicate that parental sexual orientation is not a strong predictor of individual or family outcomes. Other factors, such as the qualities of parenting and family relationships, as well as prevailing attitudes and laws in a family's environment, seem to be more important.

Behavioral adjustment has been a topic of great interest in studies of child outcomes in families with lesbian or gay parents. Early studies (e.g., Golombok et al., 1983) focused on internalizing or externalizing behavior problems as a function of parental sexual orientation, and found children of lesbian mothers to be generally well adjusted. Golombok and her colleagues (2018, 2014) studied lesbian, gay, and heterosexual parent families in the U.K., and reported that children of lesbian and gay parents were actually less likely than those of heterosexual parents to show externalizing behavior problems. Similarly, in their longitudinal study, Gartrell and her colleagues (Gartrell et al., 2000; Gartrell, Deck, Rodas, Peyser, & Banks, 2005) reported that both 5- and 10-year-old children with lesbian mothers showed comparable behavior problems as would be predicted by national norms for children with heterosexual parents. These and related data revealed that children with lesbian and gay parents develop well, with behavioral adjustment that was at least on par with that of children with heterosexual parents (Chan, Raboy, & Patterson, 1998; Goldberg & Smith, 2013).

Research has also examined children's gender development over time in families headed by lesbian, gay, and heterosexual parents. Several early studies reported no differences in children's gender development as a function of parental sexual orientation (Bos & Sandfort, 2010; Brewaeys, Ponjaert, Van Hall, & Golombok, 1997; Fulcher, Sutfin, & Patterson, 2008; Golombok et al., 2003). More recently, among adoptive families with lesbian, gay, and heterosexual parents, no significant differences were found in parents' reports or observational data on preschoolers' gender development, as a function of parental sexual orientation; across family types, children showed preferences for toys and activities typical of their gender (Farr, Bruun, Doss, & Patterson, 2018; Farr, Forssell, & Patterson, 2010a; Sumontha, Farr, & Patterson, 2017). Moreover, these findings were consistent over time—child and parent reports, in addition to observational data from early to middle childhood revealed that children's behavior was generally gender-typical and that gender development was similar across family types (Farr et al., 2018). In another study, Goldberg and Garcia (2016) examined lesbian, gay, and heterosexual parents' reports of their children's gender-typed play behavior in early childhood. Children with lesbian mothers were less likely to demonstrate gender-typical play behavior compared to children with gay and heterosexual parents across multiple time points. This could be attributed to sexual minorities being more likely to display attitudes about gender that were more flexible (Biblarz & Stacey, 2010). Overall, however, it seems that parental sexual orientation is not a strong predictor of children's gender identity or development; rather, most of the findings suggest that factors such as parents' attitudes and behaviors may be more significant.

A few studies have examined adolescent and young adult development among those with LGBTQ+ parents. Drawing data from the National Longitudinal Study of Adolescent to Adult Health, Wainright and Patterson (Wainright, Russell, & Patterson, 2004, Wainright & Patterson, 2006, 2008) examined development at 15 years of age among matched groups of youth with lesbian or heterosexual mothers. Their results showed no differences as a function of parental sexual orientation in behavior problems, peer relations, romantic lives, or substance use. Golombok and Badger (2010) reported that youth with lesbian mothers were more likely to have started heterosexual dating earlier than others of their age group. In contrast, Gartrell, Bos, and Goldberg (2012) found that female youth with lesbian mothers were older than others at their first heterosexual sexual contact. Moreover, Gartrell, Bos, and Koh (2019) found that, at 25 years of age, women with lesbian mothers were more likely than others to report same sex attractions, sexual minority identities, and same-sex sexual experiences (Gartrell et al., 2019). Even in the Gartrell et al. (2019) sample, however, most youth with lesbian mothers identified as heterosexual.

In a study of young adults who grew up with lesbian and gay parents, Lick, Patterson, and Schmidt (2013) found that there was considerable diversity in the amount of stigma that each individual had experienced. Despite that diversity, however, there were no significant differences among them in psychological adjustment. Most participants reported normative levels of adjustment, and this was true across multiple measures. Most participants identified as heterosexual, and most reported that, as they had become adults, they had become more open about having a lesbian or gay parent (Lick et al., 2013). Regardless of their sexual identities, however, those who lived in social climates that were more supportive of LGBTQ+ issues as adults reported greater well-being (Lick, Tornello, Riskind, Schmidt, & Patterson, 2012).

How do youngsters with lesbian or gay parents actually describe their experiences? Using interview data from a small sample of racially diverse adopted adolescents, Gianino, Goldberg, and Lewis (2009) explored how adolescents disclose their adoptive status and parental sexual orientation within their friendship networks and school environments. Adolescents reported a wide variety of strategies, ranging from not disclosing to anyone to telling others openly. Some participants noted that they had felt "forced" to disclose by virtue of living in a transracial adoptive family with same-sex parents, in which family members did not show physical resemblance to one another, and others described other concerns about being open with peers about their families (Gianino et al., 2009). Overall, adolescents indicated that they had received mainly positive reactions and responses from others about both their adoptive status and their lesbian or gay parents. Many other investigators have also found that there is no one single way in which children manage disclosure of parental sexual orientation; instead, youngsters employ a diverse array of strategies (e.g., Cody, Farr, McRoy, Ayers-Lopez, & Ledesma, 2017; Gershon, Tschann, & Jemerin, 1999). Interestingly, Gershon et al. (1999) also found that adolescents who were most open with peers also reported higher self-esteem. The correlation between openness and self-esteem in this study raises the question of causality: Is it that confident youngsters who feel good about themselves are more likely to disclose to peers, on one hand, or is it rather that disclosing to peers, or some third factor—as yet not identified—results in children feeling better about themselves? Research has not yet untangled possible answers to this question.

Many children and adolescents with lesbian or gay parents report having been teased or bullied by peers about parental sexual orientation (Bos & Gartrell, 2010; Bos & van Balen, 2008; Farr, Crain, Oakley, Cashen,

& Garber, 2016; Farr, Oakley, & Ollen, 2016; Gershon et al., 1999). For instance, in one study, about 40% of 17-year-old children reported having experienced teasing, bullying, and/or other forms of stigmatization at some point in their lives (Bos & Gartrell, 2010). This may be more common in some environments than others. For example, Perrin and her colleagues (2016) found that gay fathers living in a more conservative state (Tennessee) reported greater stigmatization than did those living in more liberal state (California) (Perrin et al., 2016). In the U.K., Tasker and Golombok (1997) reported that stigmatization was more likely among families who were of lower socioeconomic status.

Despite experiences of teasing or bullying, however, youth generally described positive feelings about their parents (Farr, Crain, et al., 2016). Some do report that they have experienced problems (e.g., Koh, Bos & Gartrell, 2019). It is worth noting, however, that both instances of stigma and any associated adjustment problems appear most often to be relatively mild. For example, Koh and colleagues (2019) reported that, in their study, the forms of stigma reported by most participants were "annoying questions" and "jokes." Despite their diverse experiences of stigma while growing up, most young adults with lesbian or gay parents seem to show good adjustment (e.g., Lick et al., 2013).

Other studies have focused on views of children and youth about the positive aspects of having lesbian or gay parents. Some youth describe benefits that stem from the greater motivation for parenthood that some lesbian or gay parents may be seen as having had. As one boy said, "… if you are a child of a gay or lesbian, you have a better chance of having a great parent. If you are a lesbian, you have to go through a lot of trouble to get a child, so that child is really wanted" (Cade, 1990). Others have found that adolescents with lesbian or gay parents reported feeling more open-minded and tolerant of differences displayed by others because of their experiences with parental sexual orientation (e.g., Cody et al., 2017). Potential benefits of growing up with sexual minority parents appear to be ripe for further study.

Some studies of families with lesbian and gay parents have examined outcomes for parents and for couples, as well as for parent-child relationships and overall family functioning. Patterson (1996) reported good mental health in a sample of lesbian mothers, as did Chan, Raboy, and Patterson (1998). Goldberg and Smith (2011) reported good mental health and very few depressive symptoms among a sample of lesbian and gay adoptive couples. In a study focusing on the parenting experiences of gay fathers, Tornello, Farr, and Patterson (2011) found that lesbian, gay and heterosexual participants' reports of parenting stress were well within the normative range. Farr et al. (2010a) also found that lesbian, gay, and heterosexual adoptive parents in their sample of adoptive families reported relatively little parenting stress, with no significant differences as a function of family type. Moreover, studies examining parenting stress over time among samples of lesbian, gay, and heterosexual parents have found that parenting stress was not generally associated with sexual orientation (Farr, 2017b; Goldberg & Smith, 2014; Lavner et al., 2014; Von Rijn-von Gelderen et al., 2018). Lesbian, gay, and heterosexual parents have also been found not to differ in their use of effective parenting techniques, with no significant differences in effectiveness as a function of parental sexual orientation (Farr et al., 2010a; Golombok, Blake, Casey, et al., 2018). In observational data on family interaction, lesbian, gay, and heterosexual adoptive parents were found to be relatively warm and accepting with their children overall; regardless of sexual orientation, mothers acted in warmer ways with their children than did fathers (Farr & Patterson, 2013). As in most other work in this area, the majority of individuals studied here have been white and economically secure; exploration of the extent to which findings are consistent across race, ethnicity, and economic status could be an important direction for future research.

In terms of couple relationships among lesbian and gay adoptive parents, Goldberg and Smith (2009) found that lesbian and gay adoptive couples reported relatively low levels of relationship conflict. Interestingly, Goldberg, Garcia, and Manley (2018) found higher conflict among individuals who had plurisexual identities (i.e., bisexual, queer, etc.) as compared to those with monosexual identities (i.e., lesbian, gay). In terms of additional couple relationship dynamics, Farr et al. (2010a) found that adoptive parents reported high average levels of couple relationship adjustment with no significant differences across parental sexual orientation. A majority of parents reported long-term relationships with their partners or spouses, in which they felt secure and satisfied (Farr, Forssell, & Patterson, 2010b). Lesbian and gay parents in this sample also reported overall satisfaction with current divisions of childcare labor, which participants generally described as being shared by both parents in the couple—both when children were in early childhood and in middle childhood (Farr & Patterson, 2013; Sumontha et al., 2017).

When considering the extent to which couples may break up, the research findings are not yet clear (Farr, Simon, & Goldberg, 2020).

Goldberg and Garcia (2016) reported no differences in proportions of couples who divorce as a function of sexual orientation. There have, however, been reports of higher rates of relationship dissolution among lesbian than among other parenting couples (e.g., Farr, 2017a; Gartrell, Bos, Peyser, Deck, & Rodas, 2011; MacCallum & Golombok, 2004). For example, Farr (2017a) reported that 30% of lesbian couples in her sample had separated over a period of several years, a larger number than in either of the other two groups. These reports come from relatively small samples, however, and they were recorded during a time of great legal and social change. Thus, while lesbian couple relationships may be at higher risk of dissolution, continued attention to relationships among LGBTQ+ parents seems to be warranted (Farr, Simon, & Goldberg, 2020; Farr, Vazquez, & Patterson, 2020).

Consistent with findings from the broader literature (cf. Golombok, 2015; Patterson, 2017), quality of parenting and of parent-child relationships has been more influential than parental sexual orientation in determining individual outcomes. Many studies have reported this finding (e.g., Chan, Raboy, & Patterson, 1998; Erich, Kanenberg, Case, Allen, & Bogdanos, 2009; Golombok et al., 2014; Golombok, Blake, Slutsky, et al., 2018; Wainright et al., 2004). For instance, in their study of families headed by lesbian, gay, and heterosexual adoptive couples in the U.S., Farr et al. (2010a) found that qualities of family interactions were more strongly associated with child outcomes than was family structure. Across all families, positive parenting, harmonious couple relationships, and healthy family functioning were associated with parents' reports of fewer child behavior problems when children were in early childhood and also later on, in middle childhood (Farr, 2017b; Farr et al., 2010a). Drawing on data from the same sample, Farr and Patterson (2013) found that quality of coparenting interaction was related to children's behavioral adjustment, such that more supportive and less undermining behavior between parents was associated with fewer child behavior problems. Thus, associations between parental sexual orientation and child, parent, or family outcomes have consistently been weaker than those between family processes and these outcomes.

Little is yet known about the children of transgender parents. The extant research findings are based mostly on a few small samples, but have suggested that most transgender parents tell their children directly about their transition, and that this news is received in a neutral or positive way by children (Veldorale-Griffin, 2014). When parents transition early in the children's lives, one study found that children seem to show better adjustment (White & Ettner, 2007) and another found no difference in children's adjustment (Imrie, Zadeh, Wylie, & Golombok, 2021). Overall, children and their transgender parents have been found to have strong relationships, and children have been found to show positive adjustment (Imrie et al., 2021). Negative responses by children to a parent's transition do occur, however, and if they are going to emerge, they seem to come from adolescent or adult children rather than younger ones (Grant et al., 2011). Clearly, more research with transgender parents and their children would be helpful.

3. International perspectives

For many years, most research on LGBTQ+ parents and their children was conducted in the U.S., U.K., and other English-speaking countries, but there has been a recent increase in research from other nations (Costa & Shenkman, 2020; Patterson, Riskind, & Tornello, 2014). The largest amount of this work has come from Europe—from France (e.g., Gross, 2009, Gross & Richardot, 2020), Belgium (e.g., Brewaeys et al., 1997), and the Netherlands (e.g., Bos et al., 2007). Research has also, however, emerged from Asia (e.g., Brainer, 2019, 2021), the Middle East (e.g., Erez & Shenkman, 2016; Shenkman & Shmotkin, 2014), Africa (e.g., Breshears & Lubbe-DeBeers, 2014, 2016), and Latin America (e.g., Salinas-Quiroz et al., 2018). A small number of studies have compared the experiences and adaptation of LGBTQ+ parents and their children across different national boundaries (e.g., Costa & Salinas-Quiroz, 2019; Shenkman, Gato, Tasker, Erez, & Leal, 2021).

As has been true in the U.S. and in the U.K., the diverse and rapidly changing social, legal, and cultural contexts in which this research has been conducted have proven to be intertwined with the nature of the research itself. For instance, Costa and Shenkman (2020) observed that environments in which sexual minority identities are extremely stigmatized are not conducive to completion of large-scale quantitative research projects. Perhaps for this reason, much of the research emerging from Africa and the Middle East employs qualitative methodologies with small samples of participants. Further, the nature of sexual identities themselves may vary across contexts. For these and related reasons, comparisons of findings across national borders can often be difficult to make.

When comparisons of findings across national boundaries have been possible, however, they have generally supported earlier findings. For example, Shechner, Slone, Lobel, and Shechner (2013) compared children of single and coupled lesbian versus heterosexual mothers in Israel, and found that children of single mothers showed more externalizing behavior problems than others, but maternal sexual orientation had no negative effects. Indeed, children of lesbian mothers reported more prosocial behavior and less loneliness than did those with heterosexual mothers. Thus, consistent with earlier findings, children with lesbian mothers were at least as well adjusted as those with heterosexual parents (Shechner et al., 2013).

In some cases, comparisons across national boundaries have allowed a glimpse of cultural factors that might otherwise have escaped notice. In a recent study, Shenkman and his colleagues (Shenkman et al., 2021) examined the desire for parenthood as a function of sexual orientation among childless young adults in three different nations. Two of the countries—Israel and Portugal—were considered by the authors to be strongly pronatalist, while the third—the U.K.—was described as being more characterized by individualistic values. Consistent with the pronatalist contexts in which they lived, results showed that Israeli and Portuguese participants expressed greater desire for parenthood than did their peers from the U.K., who lived in a less pronatalist context. Across all three national contexts, however, heterosexual participants expressed greater desire for parenthood than did LGBTQ+ participants who lived in the same country. Thus, cultural as well as personal characteristics were strongly associated with young adults' views about family formation.

Another approach to the understanding of social and cultural factors in this area is to study individuals who are exposed to more than one set of such values. In a recent study of this kind, Li and Patterson (2022) examined views of the future, including views about parenthood, among Chinese students who were studying as international students in the U.S. Consistent with the findings of other research, sexual minority students were less likely than heterosexual students to report a desire for parenthood. For all students, however, the perceived impact of Confucian teaching was positively associated with parenting aspirations, regardless of sexual orientation. Also independent of sexual orientation, those who described themselves as more affected by American ideas about individualism were less likely to report aspirations for parenthood. Thus, cultural ideas as well as personal characteristics were significantly associated with students' views about family formation (Li & Patterson, 2022).

Much remains to be learned about the ways in which sexual identities and experiences are shaped by law, custom, and culture, both within and across national boundaries. Attitudes and laws about sexual orientation have been changing around the world, most often—but not always—in the direction of liberalization (Flores & Park, 2018). In some nations, conservative and even reactionary ideas have gained traction in recent years, yielding even more unfavorable climates for members of sexual and gender minorities (Flores & Park, 2018). As has been the case in the U.S. and other western countries, future research and activism relevant to sexual and gender minorities seem likely to be intertwined with global social, cultural, and legal change. In describing and understanding such changes, there is much for social scientists to do.

4. Summary, conclusions, and future directions

In the U.S., many LGBTQ+ adults are parents, and many more want to have children. In their efforts to become parents, LGBTQ+ adults have reported facing numerous obstacles and experiencing many kinds of discrimination. Significant changes in the cultural and legal context in recent years, such as the legalization of marriages for same-sex couples in the U.S., have, however, placed parenthood within the realm of possibility for ever larger numbers of LGBTQ+ people. Research has revealed that, having overcome obstacles to parenthood, LGBTQ+ parents are at least as capable and effective as are heterosexual parents. Indeed, children of LGBTQ+ parents have been found to develop in ways that are similar to development among children with other parents. Quality of parenting and quality of family relationships—not parental sexual orientation or gender identity—have emerged as the best predictors of children's adjustment. Thus, as in other types of households, family processes appear to be far more central than family structure to child outcomes and to overall family functioning among LGBTQ+-parent families (Golombok, 2015; Lamb, 2012; Patterson, 2017). Early research on these topics was conducted largely in the English-speaking world, but replications of the principal findings—as well as some new insights—have also emerged from many other corners of the globe.

It is important to recognize that the impact of research findings being brought to bear on popular opinion as well as on legal decision-making has been associated with real changes in the social and legal context of LGBTQ+ people's lives, particularly in the U.S. (Gilfoyle & Dvoskin, 2017). Whereas LGBTQ+ people were once disadvantaged in many ways, LGBTQ+ people are now able, as equal citizens, to share in the legal and economic protections of marriage. This in turn has begun to make it more possible for LGBTQ+ people to consider parenthood. Moreover, research has revealed that, when LGBTQ+ people do become parents, their children generally thrive. LGBTQ+ people still show less inclination than do others to become parents (Patterson & Riskind, 2010), but the gap between heterosexual and sexual minority people in this regard may be shrinking, both in the U.S. and in many other nations of the world (Costa & Shenkman, 2020).

In view of contemporary concerns about the replicability of psychological research (e.g., Nosek & Errington, 2020), it is also valuable to acknowledge the degree to which findings on this subject have been replicated. Major findings—for example, about the successful development of children with lesbian and gay parents—have been replicated multiple times, by many investigators, using different methods, diverse samples of participants, and in many parts of the world. Results that emerged initially from research conducted in the U.S. and the U.K. have now also been reported by researchers in Belgium, Italy, the Netherlands, and other nations. This offers reason to believe that the basic findings are sound.

The robust quality of findings in this field has made them particularly valuable in legal contexts. Major findings about LGBTQ+ parents and their children have been reported and replicated many times, and they have been brought to the attention of courts via expert witness testimony and in amicus briefs, such as those offered by the American Psychological Association (Gilfoyle & Dvoskin, 2017). In this way, findings of research on LGBTQ+ parents and their children have been made available in courtrooms across the country. More recently, the research evidence has also been important in informing policy changes in many other nations.

The growth of international research on lesbian and gay parents and their children has not only resulted in important replications of major findings, but has also engendered renewed appreciation of the importance of social and cultural issues (e.g., Brainer, 2021). In a recent study conducted in Israel, Portugal, and the U.K., for example, the role of dominant ideologies was made visible in that both LGBTQ+ and heterosexual adults in Israel and Portugal were more likely to desire parenthood than were their counterparts in the U.K. (Shenkman et al., 2021). The importance of expectations about the ways in which gay parent families might be received in particular social contexts has also been documented among prospective gay fathers in Italy (Baiocco & Laghi, 2013). Further research that focuses on the social and cultural conditions in which LGBTQ+ prospective parents are living can make important contributions to knowledge in this field.

An important conclusion to draw from this body of research is the remarkable resilience that has been shown over the years by many LGBTQ+ parents and their children. Even with challenges from discriminatory treatment, many prospective LGBTQ+ parents have found ways to form families of their own. Even in the face of persistent heteronormative doubts, questions, and disapproval, many LGBTQ+ parents and their children have found ways to thrive. The ability to create families is, after all, fundamental human right, and this research has shown how, despite obstacles, LGBTQ+ people have increasingly been able to claim it.

Despite the increasing amount of research in this area, many issues remain relatively unstudied. Transgender parenting is an area in real need of study (Hafford-Letchfield et al., 2019). Diversity among LGBTQ+ parents is another such issue (Reczek, 2020). It will be important for future research to study racial, ethnic, cultural, and economic issues that affect the qualities of lives among LGBTQ+ adults and their children around the world (Brainer, Moore, & Banerjee, 2020; Goldberg, Schneebaum, Durso, & Badgett, 2020). Sexual and gender minority individuals name themselves in an increasing variety of ways (e.g., pansexual, demisexual, polyamorous, non-binary, etc.) and research is needed to understand the impact, if any, of these new identities on experiences among families (Manley & Ross, 2020). Attention should be paid to the real strengths as well as the definite challenges of these families (Coontz, 2020). Future studies would also do well to employ a variety of research designs and methodological approaches to collect data from multiple sources (Fish & Russell, 2018). In these ways, a more comprehensive understanding of LGBTQ+ parenting can emerge.

As research continues, it will be important to keep bringing its findings to the attention of those who shape the laws and policies under which LGBTQ+ people and their children live. Significant issues that loom in the U.S. may include various aspects of child welfare law and policy, such as those that affect adoption and foster care (Patterson et al., 2021; Patterson & Farr, 2022), as well as issues related to assisted reproductive technology, such as those that influence the availability of surrogacy and other forms of family formation (Cahn, 2012). These and other issues (such as legal recognition for same-sex marriage) are also likely to emerge in other nations in the coming years, and it will be helpful if empirical data from social science research can be brought to bear on them.

In conclusion, sexual and gender minority parents are growing both in numbers and in visibility around the world. Researchers have documented both some of the challenges endured by and some of the successes achieved by LGBTQ+ parents and their children. Indeed, research has shown LGBTQ+ parents and their children to be remarkably resilient against many challenges posed by heteronormativity and related ideas. Much remains to be learned, however, about the many ways in which the lives of LGBTQ+ parents and their children are shaped by the differences among them and by the qualities of the environments in which they live. In the end, such knowledge should help to enable the creation of contexts that allow more and more of these and other families to thrive.

References

Baccara, M., Collard-Wexler, A., Felli, L., & Yariv, L. (2014). Child-adoption matching: Preferences for gender and race. *American Economic Journal: Applied Economics*, *6*(3), 133–158. https://doi.org/10.1257/app.6.3.133.

Baiocco, R., & Laghi, F. (2013). Sexual orientation and the desires and intentions to become parents. *Journal of Family Studies*, *19*, 90–98.

Ball, C. A. (2012). *The right to be parents: LGBT families and the transformation of parenthood*. New York: New York University Press.

Baumrind, D. (1995). Commentary on sexual orientation: Research and policy implications. *Developmental Psychology*, *31*, 130–136.

Bergman, K., Rubio, R. J., Green, R.-J., & Padron, E. (2010). Gay men who became fathers via surrogacy: The transition to parenthood. *Journal of GLBT Family Studies*, *6*, 111–141. https://doi.org/10.1080/15504281003704942.

Berkowitz, D. (2020). Gay men and surrogacy. In A. E. Goldberg, & K. R. Allen (Eds.), *LGBTQ-parent families: Innovations in research and implications for practice* (2nd ed.). New York: Springer.

Biblarz, T. J., & Stacey, J. (2010). How does the gender of parents matter? *Journal of Marriage and Family*, *72*, 3–22. https://doi.org/10.1111/j.1741-3737.2009.00678.x.

Bos, H., & Gartrell, N. (2010). Adolescents of the USA National Longitudinal Lesbian Family Study: Can family characteristics counteract the negative effects of stigmatization? *Family Process*, *49*, 559–572.

Bos, H., & Gartrell, N. (2020). Lesbian mother families formed through donor insemination. In A. E. Goldberg, & K. R. Allen (Eds.), *LGBTQ-parent families: Innovations in research and implications for practice* (2nd ed.). New York: Springer.

Bos, H., & Sandfort, T. (2010). Children's gender identity in lesbian and heterosexual two-parent families. *Sex Roles*, *62*, 114–126.

Bos, H., & van Balen, F. (2008). Children in planned lesbian families: Stigmatization, psychological adjustment and protective factors. *Culture, Health & Sexuality*, *10*, 221–236.

Bos, H., van Balen, F., & van den Boom, D. (2007). Child adjustment and parenting in planned lesbian-parent families. *American Journal of Orthopsychiatry*, 77, 38–48. https://doi.org/10.1037/0002-9432.77.1.38.

Brainer, A. (2019). *Queer kinship and family change in Taiwan*. New Bruswick, NJ: Rutgers University Press.

Brainer, A. (2021). Lesbian and gay parents, heterosexual kinship, and queer dreams: Making families in twenty-first century Taiwan. *Positions: Asia Critique*, *29*, 633–656. https://doi.org/10.1215/10679847-8978386.

Brainer, A., Moore, M., & Banerjee, P. (2020). Race and ethnicity in the lives of LGBTQ parentsand their children: Perspectives from and beyond North America. In A. Goldberg, & K. Allen (Eds.), *LGBTQ-parent families: Innovations in research and implications for practice* (2nd ed.). New York: Springer.

Breshears, D., & Lubbe-DeBeers, C. (2014). A qualitative analysis of adult children's *advice for parents coming out to their children*. *Professional Psychology: Research and Practice*, *45*, 231–238. https://doi.org/10.1037/a0035520.

Breshears, D., & Lubbe-DeBeers, C. (2016). Same-sex parented families' negotiation of minority social identity in South Africa. *Journal of GLBT Family Studies*, *12*, 346–364. https://doi.org/10.1080/1550428X.2015k.1080134.

Brewaeys, A., Ponjaert, I., Van Hall, E. V., & Golombok, S. (1997). Donor insemination: Child development and family functioning in lesbian mother families. *Human Reproduction*, *12*, 1349–1359.

Brodzinsky, D. M. (2011). Adoption by lesbians and gay men: A national survey of adoption agency policies and practices. In D. Brodzinsky, & A. Pertman (Eds.), *Adoption by lesbians and gay men: A new dimension in family diversity* (pp. 62–84). New York, NY: Oxford University Press.

Brodzinsky, D. M., & Goldberg, A. E. (2016). Contact with birth family in adoptive families headed by lesbian, gay male, and heterosexual parents. *Children and Youth Services Review*, *62*, 9–17. https://doi.org/10.1016/j.childyouth.2016.01.014.

Brown, S., Smalling, S., Groza, V., & Ryan, S. (2009). The experiences of gay men and lesbians in becoming and being adoptive parents. *Adoption Quarterly*, *12*, 226–246. https://doi.org/10.1080/10926750903313294.

Cade, C. (1990). Two moms, no hamburgers! In L. Rafkin (Ed.), *Different mothers: Sons and daughters of lesbians talk about their lives*. Pittsburgh: Cleis Press.

Cahn, N. R. (2012). *The new kinship: Constructing donor-conceived families*. New York: NYU Press.

Chan, R. W., Brooks, R. C., Raboy, B., & Patterson, C. J. (1998). Division of labor among lesbian and heterosexual parents: Associations with children's adjustment. *Journal of Family Psychology*, *12*, 402–419.

Chan, R. W., Raboy, B., & Patterson, C. J. (1998). Psychosocial adjustment among children conceived via donor insemination by lesbian and heterosexual mothers. *Child Development*, *69*, 443–457.

Cody, P. A., Farr, R. H., McRoy, R. G., Ayers-Lopez, S. J., & Ledesma, K. J. (2017). Youth perspectives on being adopted from foster care by lesbian and gay parents: Implications for families and adoption professionals. *Adoption Quarterly*, *20*, 98–118. https://doi.org/10.1080/10926755.2016.1200702.

Coontz, S. (2020). *How to make your marriage gayer* (p. 2020). New York Times.

Costa, P. A., & Salinas-Quiroz, F. (2019). A comparative study of attitudes toward same-gender parenting and gay and lesbian rights in Portugal and Mexico. *Journal of Homosexuality*, *66*, 1909–1926.

Costa, P. A., & Shenkman, G. (2020). LGBTQ parent families in non-Western contexts. In A. E. Goldberg, & K. R. Allen (Eds.), *LGBTQ parent families: Innovations in research and implications for practice* (2nd ed.). New York Springer.

Cowan, C. P., & Cowan, P. A. (1992). *When partners become parents: The big life change for couples*. New York: Basic Books.

Dickey, L. M., Ducheny, K. M., & Ehrbar, R. D. (2016). Family creation options for transgender and gender nonconforming people. *Psychology of Sexual Orientation and Gender Diversity*, *3*, 173–179. https://doi.org/10.1037/sgd0000178.

Erez, C., & Shenkman, G. (2016). Gay dads are happier: Subjective well-being among gay and heterosexual fathers. *Journal of GLBT Family Studies*, *12*, 451–467. https://doi.org/10.1080/1550428X.2015.1102688.

Erich, S., Kanenberg, H., Case, K., Allen, T., & Bogdanos, T. (2009). An empirical analysis of factors affecting adolescent attachment in adoptive families with homosexual and straight parents. *Children & Youth Services Review*, *31*, 398–404. https://doi.org/10.1016/j.childyouth.2008.09.004.

Falk, P. J. (1989). Lesbian mothers: Psychosocial assumptions in family law. *American Psychologist*, *44*, 941–947.
Farr, R. H. (2017a). Factors associated with relationship dissolution and post-dissoluion adjustment among lesbian adoptive couples. *Journal of Lesbian Studies*, *21*, 88–105. https://doi.org/10.1080/10894160.2016.1142354.
Farr, R. H. (2017b). Does parental sexual orientation matter? A longitudinal follow-up of adoptive families with school-age children. *Developmental Psychology*, *53*, 252–264. https://doi.org/10.1037/dev0000228.
Farr, R. H., Bruun, S. T., Doss, K. M., & Patterson, C. J. (2018). Children's gender-typed behavior from early to middle childhood in adoptive families with lesbian, gay, and heterosexual parents. *Sex Roles*, *78*, 528–541. https://doi.org/10.1007/s11199-017-0812-5.
Farr, R. H., Crain, E. E., Oakley, M. K., Cashen, K. K., & Garber, K. J. (2016). Microaggressions, feelings of difference, and resilience among adopted children with sexual minority parents. *Journal of Youth and Adolescence*, *45*, 85–104. https://doi.org/10.1007/s10964-015-0353-6.
Farr, R. H., Forssell, S. L., & Patterson, C. J. (2010a). Parenting and child development in adoptive families: Does parental sexual orientation matter? *Applied Developmental Science*, *14*, 164–178. https://doi.org/10.1080/10888691.2010.500958.
Farr, R. H., Forssell, S. L., & Patterson, C. J. (2010b). Lesbian, gay, and heterosexual adoptive parents: Couple and relationship issues. *Journal of GLBT Family Studies*, *6*, 199–213. https://doi.org/10.1080/15504281003705436.
Farr, R. H., & Goldberg, A. E. (2015). Contact between birth and adoptive families during the first year post-placement: Perspectives of lesbian, gay, and heterosexual parents. *Adoption Quarterly*, *18*, 1–24. https://doi.org/10.1080/10926755.2014.895466.
Farr, R. H., & Goldberg, A. E. (2018). Sexual orientation, gender identity, and adoption law. *Family Court Review*, *56*, 374–383. https://doi.org/10.1111/fcre.12354.
Farr, R. H., Oakley, M. K., & Ollen, E. W. (2016). School experiences of young children and their lesbian and gay adoptive parents. *Psychology of Sexual Orientation and Gender Diversity*, *3*, 442–447. https://doi.org/10.1037/sgd0000187.
Farr, R. H., & Patterson, C. J. (2009). Transracial adoption by lesbian, gay, and heterosexual parents: Who completes transracial adoptions and with what results? *Adoption Quarterly*, *12*, 187–204. https://doi.org/10.1080/10926750903313328.
Farr, R. H., & Patterson, C. J. (2013). Coparenting among lesbian, gay, and heterosexual couples: Associations with adopted children's outcomes. *Child Development*, *84*, 1226–1240. https://doi.org/10.1111/cdev.12046.
Farr, R. H., Simon, K. A., & Goldberg, A. E. (2020). Separation and divorce among LGBTQ-parent families. In A. E. Goldberg, & K. R. Allen (Eds.), *LGBTQ-parent families: Innovations in research and implications for practice* (2nd ed.). New York: Springer.
Farr, R. H., Vazquez, C. P., & Patterson, C. J. (2020). LGBTQ adoptive parents and their children. In A. E. Goldberg, & K. R. Allen (Eds.), *LGBTQ-parent families: Innovations in research and implications for practice* (2nd ed.). New York: Springer.
Fish, J. N., & Russell, S. T. (2018). Queering methodologies to understand queer families. *Family Relations*, *67*, 12–25. https://doi.org/10.1111/fare.12297.
Flores, A. R., & Park, A. (2018). *Polarized progress: Social acceptance of LGBT people in 141 countries, 1981 to 2014*. The Williams Institute. Retrieved from: http://williamsinstitute.law.ucla.edu/wp-content/uploads/polarizedprogress.
Fulcher, M., Sutfin, E. L., & Patterson, C. J. (2008). Individual differences in gender development: Associations with parental sexual orientation, attitudes, and division of labor. *Sex Roles*, *58*, 330–341.
Fulton v. Philadelphia, 593 U.S. ___ (2021).
Gartrell, N., Banks, A., Reed, N., Hamilton, J., Rodes, C., & Deck, A. (2000). The National Longitudinal Lesbian Family Study: 3. Interviews with mothers of five year olds. *American Journal of Orthopsychiatry*, *70*, 542–548.

Gartrell, N., Bos, H., & Goldberg, N. G. (2012). New trends in same-sex contact for American adolescents? *Archives of Sexual Behavior*, *41*, 5–7. https://doi.org/10.1007/s10508-011-9883-5.

Gartrell, N., Bos, H., & Koh, A. (2019). National Longitudinal Lesbian Family Study – Mental health of adult offspring. *New England Journal of Medicine*, *379*, 297–299. https://doi.org/10.1056/NEJMc1804810.

Gartrell, N., Bos, H., Peyser, H., Deck, A., & Rodas, C. (2011). Family characteristics, custody arrangements, and adolescent well-being after lesbian mother bead. *Family Relations*, *60*, 572–585.

Gartrell, N., Deck, A., Rodas, C., Peyser, H., & Banks, A. (2005). The National Lesbian Family Study: 4. Interviews with the 10-year-old children. *American Journal of Orthopsychiatry*, *75*, 518–524. https://doi.org/10.1037/0002-9432.75.4.518.

Gates, G. J. (2013). *LGBT parenting in the United States*. The Williams Institute. Retrieved from http://williamsinstitute.law.ucla.edu/wp-content/uploads/LGBT-Parenting.pdf.

Gershon, T. D., Tschann, J. M., & Jemerin, J. M. (1999). Stigmatization, self-esteem, and coping among the adolescent children of lesbian mothers. *Journal of Adolescent Health*, *24*, 437–445.

Gianino, M., Goldberg, A., & Lewis, T. (2009). Disclosure practices among adopted youth with gay and lesbian parents. *Adoption Quarterly*, *12*, 205–228. https://doi.org/10.1080/10926750903313344.

Gilfoyle, N., & Dvoskin, J. A. (2017). APA's amicus curiae program: Bringing psychological research to judicial decisions. *American Psychologist*, *72*, 753–763. https://doi.org/10.1037/amp0000221.

Goldberg, A. E. (2009a). Lesbian and heterosexual preadoptive couples' openness to transracial adoption. *American Journal of Orthopsychiatry*, *79*, 103–117. https://doi.org/10.1037/a0015354.

Goldberg, A. E. (2009b). Heterosexual, lesbian, and gay preadoptive parents' preferences about child gender. *Sex Roles*, *61*, 55–71. https://doi.org/10.1007/s11199-009-9598-4.

Goldberg, A. E. (2012). *Gay dads: Transitions to adoptive fatherhood*. New York, NY: NYU Press.

Goldberg, A. E., & Allen, K. R. (Eds.). (2020). *LGBTQ-parent families: Innovations in research and implications for practice* (2nd ed.). New York: Springer.

Goldberg, S. K., & Conron, K. J. (2018). *How many same-sex couples in the U.S. are raising children?*. The Williams Institute. Retrieved from https://williamsinstitute.law.ucla.edu/research/parenting/how-many-same-sex-parents-in-us/.

Goldberg, A. E., Downing, J. B., & Richardson, H. B. (2009). The transition from infertility to adoption: Perceptions of lesbian and heterosexual preadoptive couples. *Journal of Social & Personal Relationships*, *26*, 938–963. https://doi.org/10.1177/0265407509345652.

Goldberg, A. E., & Garcia, R. L. (2016). Gender-typed behavior over time in children with lesbian, gay, and heterosexual parents. *Journal of Family Psychology*, *30*, 854–865. https://doi.org/10.1037/fam0000226.

Goldberg, A. E., Garcia, R., & Manley, M. H. (2018). Monosexual and nonmonosexual women in same-sex couples' relationship quality during the first five years of parenthood. *Sexual and Relationship Therapy*, *33*, 190–212. https://doi.org/10.1080/14681994.2017.1419561.

Goldberg, A. E., Gartrell, N. K., & Gates, G. J. (2014). *Research report on LGB-parent families*. The Williams Institute. Retrieved from https://escholarship.org/uc/item/7gr4970w.

Goldberg, A. E., Kinkler, L. A., Moyer, A. M., & Weber, E. (2014). Intimate relationship challenges in early parenthood among lesbian, gay, and heterosexual couples adopting via the child welfare system. *Professional Psychology*, *45*, 221–230. https://doi.org/10.1037/a0037443.

Goldberg, A. E., Kinkler, L. A., Richardson, H. B., & Downing, J. B. (2011). Lesbian, gay, and heterosexual couples in open adoption arrangements: A qualitative study. *Journal of Marriage and Family*, *73*, 502–518. https://doi.org/10.1111/j.1741-3737.2010.00821.x.

Goldberg, A. E., Moyer, A. M., Kinkler, L. A., & Richardson, H. B. (2012). "When you're sitting on the fence, hope's the hardest part": Challenges and experiences of heterosexual and same-sex couples adopting through the child welfare system. *Adoption Quarterly*, *15*, 288–315. https://doi.org/10.1080/10926755.2012.731032.

Goldberg, N. G., Schneebaum, A., Durso, L. E., & Badgett, M. V. L. (2020). LGBTQ parent families in the United States and economic well-being. In A. Goldberg, & K. Allen (Eds.), *LGBTQ-parent families: Innovations in research and implications for practice* (2nd ed.). New York: Springer.

Goldberg, A. E., & Smith, J. Z. (2008). Social support and psychological well-being in lesbian and heterosexual preadoptive couples. *Family Relations*, *57*, 281–294. https://doi.org/10.1111/j.1741-3729.2008.00500.x.

Goldberg, A. E., & Smith, J. Z. (2009). Perceived parenting skill across the transition to adoptive parenthood among lesbian, gay, and heterosexual couples. *Journal of Family Psychology*, *23*, 861–870. https://doi.org/10.1037/a0017009.

Goldberg, A. E., & Smith, J. Z. (2011). Stigma, social context, and mental health: Lesbian and gay couples across the transition to adoptive parenthood. *Journal of Counseling Psychology*, *58*, 139–150. https://doi.org/10.1037/a0021684.

Goldberg, A. E., & Smith, J. Z. (2013). Predictors of psychological adjustment in early placed adopted children with lesbian, gay, and heterosexual parents. *Journal of Family Psychology*, *27*, 431–442. https://doi.org/10.1037/a0032911.

Goldberg, A. E., & Smith, J. Z. (2014). Predictors of parenting stress in lesbian, gay, and heterosexual adoptive parents during early parenthood. *Journal of Family Psychology*, *28*, 125–137. https://doi.org/10.1037/a0036007.

Goldberg, A. E., Smith, J. Z., & Kashy, D. A. (2010). Preadoptive factors predicting lesbian, gay, and heterosexual couples' relationship quality across the transition to adoptive parenthood. *Journal of Family Psychology*, *24*, 221–232. https://doi.org/10.1037/a0019615.

Goldberg, A. E., Tornello, S., Farr, R., Smith, J. Z., & Miranda, L. (2020). Barriers to adoption and foster care and openness to child characteristics among transgender adults. *Children and Youth Services Review*, *109*, 104699. https://doi.org/10.1016/j.childyouth.2019.104699.

Golombok, S. (2015). *Modern families: Parents and children in new family forms*. Cambridge UK: Cambridge University Press.

Golombok, S., & Badger, S. (2010). Children raised in mother-headed families from infancy: A follow-up of children of lesbian and single heterosexual mothers at early adulthood. *Human Reproduction*, *25*, 150–157. https://doi.org/10.1093/humrep/dep345.

Golombok, S., Blake, L., Casey, P., Roman, G., & Jadva, V. (2018). Children born through reproductive donation: A longitudinal study of child adjustment. *Journal of Child Psychology and Psychiatry*, *54*, 653–660. https://doi.org/10.1111/jcpp.12015.

Golombok, S., Blake, L., Slutsky, J., Raffanello, E., Roman, G., & Eherhardt, A. (2018). Parenting and the adjustment of children born to gay fathers through surrogacy. *Child Development*, *89*, 1223–1233. https://doi.org/10.1111/cdev.12728.

Golombok, S., Mellish, L., Jennings, S., Casey, P., Tasker, F., & Lamb, M. E. (2014). Adoptive gay father families: Parent-child relationships and children's psychological adjustment. *Child Development*, *85*, 456–468. https://doi.org/10.1111/cdev.12155.

Golombok, S., Perry, B., Burston, A., Murray, C., Mooney-Somers, J., Stevens, M., et al. (2003). Children with lesbian parents: A community study. *British Journal of Developmental Psychology*, *39*, 20–33. https://doi.org/10.1037/0012-1649.39.1.20.

Golombok, S., Spencer, A., & Rutter, M. (1983). Children in lesbian and single-parent households: Psychosexual and psychiatric appraisal. *Journal of Child Psychology and Psychiatry*, *24*, 551–572.

Grant, J. M., Mottet, L. A., & Tanis, J. (2011). *Injustice at every turn: A report of the National Transgender Discrimination Survey*. Washington: National Center for Transgender Equality and National Gay and Lesbian Task Force.

Gross, M. (2009). The desire for parenthood among lesbians and gay men. In D. Marre (Ed.), *International adoption*. New York: NYU Press.

Gross, M., & Richardot, S. (2020). Conception narratives and third party reproduction in lesbian-parented families in France. *International Social Science Journal*, *70*, 235–236.

Hafford-Letchfield, T., Cocker, C., Rutter, D., Tinarwo, M., McCormack, K., & Manning, R. (2019). What do we know about transgender parenting? Findings from a systematic review. *Health and Social Care in the Community*, *27*, 1111–1125.

Herrmann-Green, L. K., & Gehring, T. M. (2007). The German lesbian family study: Planning for parenthood via donor insemination. *Journal of GLBT Family Studies*, *3*, 351–395. https://doi.org/10.1300/J461v03n04_02.

Hitchens, D. J., & Kirkpatrick, M. J. (1985). Lesbian mothers/gay fathers. In D. H. Schetky, & E. P. Benedek (Eds.), *Emerging issues in child psychiatry and the law* (pp. 115–125). New York: Brunner/Mazel.

Imrie, S., Zadeh, S., Wylie, K., & Golombok, S. (2021). Children with trans parents: Parent-child relationship quality and psychological well-being. *Parenting*, *21*, 185–215. https://doi.org/10.1080/15295192.2020.1792194.

Jennings, S., Mellish, L., Tasker, F., Lamb, M., & Golombok, S. (2014). Why adoption? Gay, lesbian, and heterosexual adoptive parents' reproductive experiences and reasons for adoption. *Adoption Quarterly*, *17*(3), 205–226.

Kinkler, L. A., & Goldberg, A. E. (2011). Working with what we've got: Perceptions of barriers and supports among small-metropolitan-area same-sex adopting couples. *Family Relations*, *60*, 387–403. https://doi.org/10.1111/j.1741-3729.2011.00654.x.

Koh, A. S., Bos, H. M. S. & Gartrell, N. K. (2019). Predictors of mental health in emerging adult offspring of lesbian-parent families. *Journal of Lesbian Studies*, *23*, 257–278. https://doi.org/10.1080/10894160.2018.1555694.

Lamb, M. E. (2012). Mothers, fathers, families, and circumstances: Factors affecting children's adjustment. *Applied Developmental Science*, *16*, 98–111. https://doi.org/10.1080/10888691.2012.667344.

Lavner, J. A., Waterman, J., & Peplau, L. A. (2012). Can gay and lesbian parents promote healthy development in high-risk children adopted from foster care? *American Journal of Orthopsychiatry*, *82*, 465–472. https://doi.org/10.1111/j.1939-0025.2012.01176.x.

Lavner, J. A., Waterman, J., & Peplau, L. A. (2014). Parent adjustment over time in gay, lesbian, and heterosexual parent families adopting from foster care. *American Journal of Orthopsychiatry*, *84*(46–53). https://doi.org/10.1037/h0098853.

Lawrence v. Texas, 539 U.S. 558, 578 (2003).

Li, Y., & Patterson, C. J. (2022). Parenting aspirations among Chinese international students of diverse sexual identities: A cultural perspective. *Journal of GLBT Family Studies* (in press).

Lick, D. J., Patterson, C. J., & Schmidt, K. M. (2013). Recalled social experiences and current psychological adjustment among adults reared by gay and lesbian parents. *Journal of GLBT Family Studies*, *9*, 230–253. https://doi.org/10.1080/1550428X.2013.781907.

Lick, D. J., Tornello, S. L., Riskind, R. G., Schmidt, K. G., & Patterson, C. J. (2012). Social 'climate for sexual minorities predicts well-being among heterosexual offspring of lesbian and gay parents. *Journal of Sexuality Research and Social Policy*, *9*, 99–112. https://doi.org/10.1007/s13178-012-0081-6.

MacCallum, F., & Golombok, S. (2004). Children raised in fatherless families from infancy: A follow-up of children of lesbian and single heterosexual mothers at early adolescence. *Journal of Child Psychology and Psychiatry*, *45*, 1407–1419.

Mallon, G. P. (2011). The home study assessment process for gay, lesbian, bisexual, and transgender prospective foster and adoptive families. *Journal of GLBT Family Studies*, 7, 9–29. https://doi.org/10.1080/1550428X.2011.537229.

McKay, K., Ross, L. E., & Goldberg, A. E. (2010). Adaptation to parenthood during the post-adoption period: A review of the literature. *Adoption Quarterly*, *13*, 125–144. https://doi.org/10.1080/10926755.2010.481040.

Moore, M. R. (2011). *Invisible families: Gay identities, relationships, and motherhood among black women*. Berkeley CA: University of California Press.

Moore, M. R., & Stambolis-Ruhstorfer, M. (2013). LGBT sexuality and families at the start of the twenty-first century. *Annual Review of Sociology*, *39*, 491–507. https://doi.org/10.1146/annurev-soc-071312-145643.

Movement advancement project. (2018). *Creating a license to discriminate: 2018 Federal Child Welfare Amendment. (issue brief)*. Retrieved from http://www.lgbtmap.org/2018-child-welfare-amendment.

Nosek, B. A., & Errington, T. M. (2020). What is replication? *PLoS Biology*, *18*, e3000691. https://doi.org/10.1371/journal.pbio.3000691.

Obergefell v. Hodges, 135 S. Ct. 2584 (2015).

Patterson, C. J. (1992). Children of lesbian and gay parents. *Child Development*, *63*, 1025–1042.

Patterson, C. J. (1995a). Families of the lesbian baby boom: Parents' division of labor and children's adjustment. *Developmental Psychology*, *31*, 115–123.

Patterson, C. J. (1995b). Adoption of minor children by lesbian and gay adults: A social science perspective. *Duke Journal of Gender Law and Policy*, *2*, 191–205.

Patterson, C. J. (1996). Lesbian mothers and their children: Findings from the Bay Area families study. In J. Laird, & R. J. Green (Eds.), *Lesbians and gays in couples and families: A handbook for therapists* (pp. 420–437). San Francisco: Jossey-Bass.

Patterson, C. J. (2000). Family relationships of lesbians and gay men: A decade review. *Journal of Marriage and the Family*, *62*, 1052–1069.

Patterson, C. J. (2006). Children of lesbian and gay parents. *Current Directions in Psychological Science*, *15*, 241–244.

Patterson, C. J. (2009). Children of lesbian and gay parents: Psychology, law, and policy. *American Psychologist*, *64*, 727–736.

Patterson, C. J. (2013). Children of lesbian and gay parents: Psychology, law, and policy. *Psychology of Sexual Orientation and Gender Diversity*, *1*, 27–34. https://doi.org/10.1037/2329-0382.1.S.27.

Patterson, C. J. (2017). Parents' sexual orientation and children's development. *Child Development Perspectives*, *11*, 45–49. https://doi.org/10.1111/cdep.12207.

Patterson, C. J., & Farr, R. H. (2017). What shall we call ourselves? Last names among lesbian, gay, and heterosexual couples and their adopted children. *Journal of GLBT Family Studies*, *13*, 97–113. https://doi.org/10.1080/1550428X.2016.1169239.

Patterson, C. J., & Farr, R. H. (2022). Sexual orientation, gender identity, and Foster Care: What can social science offer in a case like *Fulton v. City of Philadelphia*? *Family Law Review*.

Patterson, C. J., Farr, R. H., & Goldberg, A. E. (2021). *LGBTQ+ parents and their children*. Minneapolis MN. Author: National Council on Family Relations Policy Brief.

Patterson, C. J., & Riskind, R. G. (2010). To be a parent: Issues in family formation among gay and lesbian adults. *Journal of GLBT Family Studies*, *6*, 326–340.

Patterson, C. J., Riskind, R. G., & Tornello, S. L. (2014). Sexual orientation and parenting: A global perspective. In A. Abela, & J. Walker (Eds.), *Contemporary issues in family studies: Global perspectives on partnerships, parenting, and support in a changing world*. New York: Wiley/Blackwell.

Patterson, C. J., Sepulveda, M. J., & White, J. (2020). *Understanding the well-being of LGBTQI + populations*. Washington, DC: The National Academies Press. https://doi.org/10.7226/25877.

Patterson, C. J., Sutfin, E. L., & Fulcher, M. (2004). Division of labor among lesbian and heterosexual parenting couples: Correlates of specialized versus shared patterns. *Journal of Adult Development*, *11*, 179–189.

Patterson, C. J., & Tornello, S. L. (2011). Gay fathers' pathways to parenthood: International perspectives. *Zeitschrift für Familienforschung (Journal of Family Research)*, 7, 103–116.

Perrin, E., Hurley, S., Mattern, K., Flavin, L., & Pinderhughes, E. (2019). Barriers and stigma experienced by gay fathers and their children. *Pediatrics*, *143*, e20180683.

Perrin, E., Pinderhughes, E., Mattern, K., Hurley, S., & Newman, R. (2016). Experiences of children with gay fathers. *Clinical Pediatrics*, *55*, 1305–1317. https://doi.org/10.1177/0009922816632346.

Pilcher, J. (2017). Names and "doing gender": How forenames and surnames contribute to gender identities, difference, and inequalities. *Sex Roles*, 77, 812–822. https://doi.org/10.1007/s11199-017-0805-4.

Pyne, J. (2012). *Transforming family: Trans parents and their struggles, strategies, and strengths*. Toronto, Ontario: LGBTQ Parenting Network, Sherbourne Health Clinic.

Raleigh, E. (2012). Are same-sex and single adoptive parents more likely to adopt transracially? A national analysis of race, family structure, and the adoption marketplace. *Sociological Perspectives*, *55*, 449–471. https://doi.org/10.1525/sop.2012.55.3.449.

Reczek, C. (2020). Sexual- and gender-minority families: A 2010 to 2020 decade in review. *Journal of Marriage and Family*, *82*, 300–325. https://doi.org/10.1111/jomf.12607.

Von Rijn-von Gelderen, L., Bos, H., Jorgenson, T. D., Ellis-Davies, K., Winstanley, A., Golombok, S., et al. (2018). Wellbeing of gay fathers with children born through surrogacy: A comparison with lesbian-mother families and heterosexual IVF parent families. *Human Reproduction*, *33*, 101–108.

Riskind, R., & Patterson, C. J. (2010). Parenting intentions and desires among childless lesbian, gay, and heterosexual individuals. *Journal of Family Psychology*, *24*, 78–81.

Riskind, R. G., & Tornello, S. L. (2017). Sexual orientation and future parenthood in a 2011–2013 nationally representative United States sample. *Journal of Family Psychology*, *31*, 792–798. https://doi.org/10.1037/fam0000316.

Rivers, D. R. (2013). *Radical relations: Lesbian mothers, gay fathers, and their children in the United States since world war II*. Chapel Hill: University of North Carolina Press.

Manley, M., & Ross, L. (2020). What do we know about bisexual parenting? A continuing call for research. In A. Goldberg, & K. Allen (Eds.), *LGBTQ-parent families: Innovations in research and implications for practice* (2nd ed.). New York: Springer.

Ross, L. E., Epstein, R., Anderson, S., & Eady, A. (2009). Policy, practice, and personal narratives: Experiences of LGBTQ people with adoption in Ontario, Canada. *Adoption Quarterly*, *12*, 272–293. https://doi.org/10.1080/10926750903313302.

Ross, L. E., Epstein, R., Goldfinger, C., Steele, L., Anderson, S., & Strike, C. (2008). Lesbian and queer mothers navigating the adoption system: The impacts on mental health. *Health Sociology Review*, *17*, 254–266.

Salinas-Quiroz, F., Rodríguez-Sánchez, F., Costa, P. A., Rosales, M., Silva, P., & Cambón, V. (2018). Can children have ordinary expectable caregiving environments in unconventional contexts? Quality of care organization in three mexican same-sex planned families. *Frontiers in Psychology*, *9*, 2349.

Shapiro, J. (2020). The law governing LGBTQ-parent families in the United States. In A. E. Goldberg, & K. R. Allen (Eds.), *LGBTQ-parent families: Innovations in research and implications for practice* (2nd ed.). New York: Springer.

Shechner, T., Slone, M., Lobel, T., & Shechner, R. (2013). Children's adjustment in non-traditional families in Israel: The effect of parental sexual orientation and the number of parents on children's development. *Child: Care, Health and Development*, *39*, 178–184. https://doi.org/10.1111/j.1365-2214.2011.01337.x.

Shenkman, G., Gato, J., Tasker, F., Erez, C., & Leal, D. (2021). Deciding to parent or remain childfree: Comparing sexual minority and heterosexual childless adults from Israel, Portugal, and the United Kingdom. *Journal of Family Psychology*. https://doi.org/10.1027/fam0000843. Advance online publication.

Shenkman, G., & Shmotkin, D. (2014). "Kids are joy": Psychological welfare among Israeil gay fathers. *Journal of Family Issues*, *35*, 1926–1939. https://doi.org/10.1177/0192513X13489300.

Simon, K. A., Tornello, S. L., Farr, R. H., & Bos, H. M. (2018). Envisioning future parenthood among bisexual, lesbian, and heterosexual women. *Psychology of Sexual Orientation and Gender Diversity*, *5*, 253–259. https://doi.org/10.1037/sgd0000267.

Stotzer, R. L., Herman, J. L., & Hasenbush, A. (2014). *Transgender parenting: A review of existing literature*. The Williams Institute: Retrieved from. https://escholarship.org/uc/item/3rp0v7qv.

Sumontha, J., Farr, R. H., & Patterson, C. J. (2016). Social support and coparenting among lesbian, gay, and heterosexual adoptive parents. *Journal of Family Psychology*, *30*, 987–996. https://doi.org/10.1037/fam0000253.

Sumontha, J., Farr, R. H., & Patterson, C. J. (2017). Children's gender development: Associations with parental sexual orientation, division of labor, and gender ideology. *Psychology of Sexual Orientation and Gender Diversity*, *4*, 438–450. https://doi.org/10.1037/sgd0000242.

Tasker, F. L., & Golombok, S. (1997). *Growing up in a lesbian family*. London: Guilford.

Tornello, S. L., & Bos, H. (2017). Parenting intentions among transgender individuals. *LGBT Health*, *4*, 115–120. https://doi.org/10.1089/lgbt.2016.0153.

Tornello, S. L., Farr, R. H., & Patterson, C. J. (2011). Predictors of parenting stress among gay adoptive fathers in the United States. *Journal of Family Psychology*, *25*, 591–600. https://doi.org/10.1037/a0024480.

Tornello, S. L., Kruczkowski, S., & Patterson, C. J. (2015). Division of labor and relationship quality among male same-sex couples who became fathers via surrogacy. *Journal of GLBT Family Studies*, *11*, 375–394.

Tornello, S. L., & Patterson, C. J. (2015). Timing of parenthood and experiences of gay fathers: A life course perspective. *Journal of GLBT Family Studies*, *11*, 35–56. https://doi.org/10.1080/1550428X.2013.878681.

Veldorale-Griffin, A. (2014). Transgender parents and their adult children's experiences of disclosure and transition. *Journal of GLBT Family Studies*, *10*, 475–501. https://doi.org/10.1080/1550428X.2013.866063.

Wainright, J. L., & Patterson, C. J. (2006). Delinquency, victimization, and substance use among adolescents with female same-sex parents. *Journal of Family Psychology*, *20*, 526–530.

Wainright, J. L., & Patterson, C. J. (2008). Peer relations among adolescents with female same-sex parents. *Developmental Psychology*, *44*, 117–126.

Wainright, J. L., Russell, S. T., & Patterson, C. J. (2004). Psychosocial adjustment and school outcomes of adolescents with same-sex parents. *Child Development*, *75*, 1886–1898.

White, T., & Ettner, R. (2007). Adaptation and adjustment in children of transsexual parents. *European Child and Adolescent Psychiatry*, *16*, 215–221. https://doi.org/10.1007/s00787-006-0591-y.

CHAPTER FOUR

Environmental influences on early language and literacy development: Social policy and educational implications

Meredith L. Rowe*
Harvard Graduate School of Education, Harvard University, Cambridge, MA, United States
*Corresponding author: e-mail address: meredith_rowe@gse.harvard.edu

Contents

Abstract

A growing body of literature suggests strong associations between environmental factors and young children's early language and literacy development. In the United States, large socio-economic differences are evident in children's skills when they enter Kindergarten, differences that persist through schooling and can be explained by children's early communicative environments. Here, I highlight three themes that characterize the features of children's communicative environments that are found to promote language learning: (1) Talking *with* children helps more than talking to children, (2) Linguistic input should increase in diversity and complexity

Advances in Child Development and Behavior, Volume 63
ISSN 0065-2407
https://doi.org/10.1016/bs.acdb.2022.04.001

during early childhood, and (3) A gradual transition from contextualized to decontextualized conversations is helpful. There are many reasons for the large variability in early communicative environments within and across social class groups. Two primary reasons include parents' knowledge of child development and parenting stress. Social policies that reduce parenting stress and increase parental knowledge have the potential to improve early language environments and lead to better educational outcomes for all children.

1. Introduction

In the United States, large socioeconomic differences in children's language and literacy skills are evident when they enter kindergarten. For example, analyses of National datasets reveal a greater than 1 standard deviation difference in reading skills at kindergarten entry between children whose family income is at the 90th percentile versus the 10th percentile (Garcia & Weiss, 2017; Reardon, 2013), suggesting that some American children start school more than a grade ahead of their same-aged peers in skill level. However, these differences by socio-economic status (SES) are *not* evident at birth. For example, studies show no socioeconomic differences in babies' brains at birth, measured as resting EEG power which is associated with later cognitive abilities (Brito, Fifer, Myers, Elliott, & Noble, 2016). Thus, the large socioeconomic gaps in skills evident at school entry, must develop during the early childhood years. Indeed, brain development during early childhood is characterized by a large amount of plasticity, highlighting the strong potential influence of the environment in early language development (Huttenlocher, 2009). Confirming this plasticity, effects of SES on brain development emerge in the first year of life (Tomalski et al., 2013) and increase with age over the early childhood period (Noble, Houston, Kan, & Sowell, 2012).

We also know that language development is a social process (e.g., Kuhl, 2007), influenced by contextual factors (e.g., Rowe & Weisleder, 2020). Variations in opportunities to engage in language-rich social interactions with others during early childhood can have large effects on early language development. In the United States context, these variations are evident, and significant, within social classes but even more pronounced across social classes (e.g., Hirsh-Pasek et al., 2015; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Rowe, 2012; Weisleder & Fernald, 2013). To address the pressing issue of inequity in early opportunities and skills, in this chapter I first summarize what we know about the attributes of early *communicative* environments that promote language learning across early childhood. I then explain two of the reasons why children's communicative environments may vary so much across families, highlighting the impact of parents' knowledge of child development and parenting stress. Finally, I discuss implications for policy and suggest several strategies that could indirectly enhance early communicative environments by reducing parenting stress and increasing parenting knowledge and mindsets. But first, I offer a quick summary of what we know about language development as a foundation for literacy skills and why a focus on early language environments can promote not just language but literacy more broadly.

2. From language to literacy

As noted above, American kindergartners enter school with widely variable reading skills, but what exactly do we mean by "reading" in kindergarten? And why are the skills that students start with so influential for later school success? Most kindergartners do not come to school able to pick up a book and read text fluently. However, they arrive with a foundation of oral language and code-based skills that will help them learn to read. The code-based skills refer to knowledge that is most likely to be explicitly taught, for example knowing the names of the letters of the alphabet and understanding the sounds that letters can make (phonics). Systematically teaching these code-based skills is an essential component of reading instruction in the early grades. The oral language skills children bring with them include vocabulary, syntax, listening comprehension, and narrative skills. In contrast to the code-based skills, these oral language skills are developing from birth and serve as a foundation for the code-based skills (Dickinson, Golinkoff, & Hirsh-Pasek, 2010; Duff, Reen, Plunkett, & Nation, 2015; Foorman, Herrera, Petscher, Mitchell, & Truckenmiller, 2015; NICHD, 2005).

Studies consistently find that oral language skills in kindergarten predict reading comprehension skills in later grades and often mediate, or explain, the effects of SES on reading comprehension (e.g., Durham, Farkas, Hammer, Tomblin, & Catts, 2007; Maguire et al., 2018). This is important because it suggests that it is not the child's SES that determines their reading ability, but the oral language skills they bring with them when they enter school. Thus, efforts to enhance the development of oral language skills during early childhood should lead to better literacy outcomes. It is not surprising that oral language skills relate to comprehension, as even once a child can decode a word, they need to know the meaning of the word to make sense of the text. Beyond vocabulary, experience with oral language that is decontextualized, or more abstract (talking with parents at home about things that happened in the past or might happen in the future) provides children with a basis to participate in decontextualized classroom discussions and to interpret texts that are also decontextualized (Snow, 1991). Indeed, we know that children who are exposed to more oral language during early childhood, are better able to process oral language and build their vocabularies (Fernald, Perfors, & Marchman, 2006; Weisleder & Fernald, 2013). Further, children who have opportunities to engage with more decontextualized language in the home environment build greater vocabulary, syntax, and narrative skills than those with fewer opportunities (Demir, Rowe, Heller, Goldin-Meadow, & Levine, 2015). With oral language skills emerging from birth, the role of the environment is crucial, as skills build upon skills and those who start ahead tend to stay ahead (Heckman, 2011).

3. Meaningful variations in early communicative environments

Over the past three decades, substantial research has contributed to our understanding of the attributes of young children's early communicative environments that contribute to their language development. Here I highlight several of the main themes that have emerged in the literature and provide some examples of each from our own work and that of others.[a]

3.1 Talking *with* children helps more than talking to children

When initial findings on the role of parent input in child vocabulary growth surfaced in the 1990s (e.g., Hart & Risley, 1995; Huttenlocher et al., 1991) the most pressing take-home message was that it is the number of words that children hear that influences language learning. Fortunately, that early work spurred much more research and we now have a more nuanced understanding of the features of the communicative environment that promote

[a] It is important to acknowledge that this review is limited to research with children growing up in Western societies with existing socioeconomic inequalities in language and literacy skills. It is quite possible that in other cultures/societies there may be different patterns of communication that lead to learning.

language development. One theme that is clear from the research to date is that children learn more when they are engaged in conversations than when they are merely exposed to words—the social-interaction is key.

To start, there is evidence that infants cannot learn the phonetic properties of a language by watching and listening to someone speak that language on a television screen, whereas the same language spoken from a person interacting with the infant promotes learning (Kuhl, 2007). Additional evidence comes from studies that examine home environments and look at relations between child-directed speech versus ambient or overheard speech and language learning. Results from various cultures suggest that in home environments child-directed speech is significantly associated with language development for toddlers, whereas the overheard or ambient speech is not (e.g., Golinkoff, Hoff, Rowe, Tamis-LeMonda, & Hirsh-Pasek, 2019; Shneidman & Goldin-Meadow, 2012; Weisleder & Fernald, 2013). Within the scope of child-directed speech, research shows an enhanced effect of speech which occurs during episodes of joint attention, where the parent and toddler are both focusing on one another and another object or event. This talk during joint episodes of attention is more strongly related to learning than talk outside of these episodes (Tomasello & Farrar, 1986), and infants are more likely to learn words when parents talk with them about the objects that they are looking at (Yu & Smith, 2012), highlighting the important role of infant attention in language learning (Shneidman & Woodward, 2016).

Similarly, engaging young children in back-and-forth conversations has larger effects on language development than just talking to children. For example, a recent study examined interaction patterns between 1-year-old infants and their mothers and found that the best predictor of later expressive language in the children were situations where the children were intentionally communicating by looking at the parent and vocalizing and the caregiver responded contingently to that communication, highlighting the importance of the integrated social interaction (Donnellan, Bannard, McGillion, Slocombe, & Matthews, 2020). Further, Hirsh-Pasek and colleagues studied a large low-income sample of American families and found that the extent to which the interaction between parents and toddlers was fluid and connected with back-and-forth turn taking was a better predictor of language than the quantity of parent words spoken during the interaction (Hirsh-Pasek et al., 2015). Thus, beginning in infancy, engaging children in social, back-and-forth, communicative interactions predicts vocabulary development.

In our work, led by Romeo et al. (2018) we explored the neural and behavioral correlates of children's exposure to conversational turns at home in a socio-economically diverse sample of families with 4–6-years-old in the Boston area. We used the LENA recording device to obtain measures of the adult words the children heard at home, the number of words children produced, and the number of conversational turns between the parent and child at home. We measured the children's verbal skills using a battery of oral language measures (vocabulary, syntax, listening comprehension) in the lab which were combined into a composite measure of verbal skill, and we measured neural language processing through a story-listening functional MRI task where the contrast of interest was the difference in activation during a forward-speech listening condition versus backward-speech listening condition. We found that the conversational turn measure predicted children's verbal skill over and above SES (education and income), but the adult words and child words measures did not. Further, children who had experienced more conversational turns with adults at home, independent of SES, IQ, and adult or child utterances alone, exhibited greater left inferior frontal (Broca's area) activation during the fMRI task, which significantly explained the relation between children's language exposure and verbal skill. This suggests that over and above SES, it was the conversational turn experience at home that predicted children's neural activation patterns which were associated with their oral language skills. The take-home message from this study and the others mentioned above is that talking *with* children across the early childhood period seems to be more effective at promoting oral language skills than merely talking to children.

3.2 Helpful input increases in diversity and complexity as children age

A second theme that emerges from the literature on parent input and child language development is the fact that as children get older and increase in language abilities themselves, the features of the linguistic input that are most helpful change. This isn't so surprising, as anyone who spends time with young children knows that you don't typically talk to a 1-year-old the same way you talk with a 4-year-old, and therefore what a 1-year-old needs to promote language learning is different from the language experience that would benefit a 4-year-old. As a rule of thumb, increasing the lexical diversity and syntactic complexity of conversations with children over the early childhood period should be helpful. This is in line with Vygotsky's sociocultural theory which emphasizes the social influences on learning and highlights how each child's zone of proximal development

(Vygotsky, 1978) changes as they increase in ability. In this case, helpful input is just a little linguistically challenging for that zone and comes with appropriate support to keep the child engaged in conversation. We provide some examples of this developmental shift in helpful features of linguistic input from our own work and others' below.

In the first year of life infants often need multiple exposures to words to learn them. In one study with parents and their 7-months-old infants, we found that it was the repetition in the parent input that was most positively related to children's vocabulary at age 2 (Newman, Rowe, & Ratner, 2016). However, by the time infants develop into toddlers they are quicker at picking up new words and repetition may not be as helpful. We tested this out in a study of 41 low-income fathers and their 2-years-old. Here, we used several different measures of lexical repetition in the fathers' speech and found that across all the measures, repetition was negatively associated with child vocabulary. That is, fathers whose children had larger vocabularies at age two, used repetition less often than fathers whose children had smaller vocabularies (Schwab, Rowe, Cabrera, & Lew-Williams, 2018). Indeed, in another study we found that at child age 30-months, the diversity of parent words used or the number of rare or sophisticated words (not the 3000 most common words) that parents used was a strong predictor of children's vocabulary growth, over and above the amount that parents talked and SES (Rowe, 2012). We found that this effect of vocabulary diversity held for both typically developing children and children who had experienced early pre- or peri-natal brain injury (see Fig. 1A). Weizman and Snow (2001)

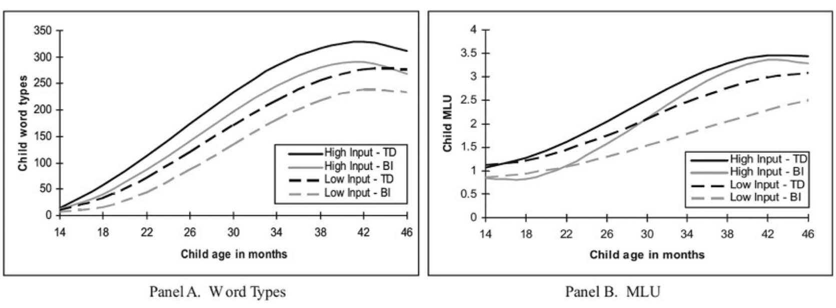


Fig. 1 Effect of parent vocabulary diversity (word types, panel A) and mean length of utterance (MLU, Panel B) at child age 30-months on change over time in children's word types and MLU for typically developing children (TD) and children who experienced brain injury (BI), controlling for parent education level ($n=80$). *Reprinted from Rowe, M. L., Levine, S. C., Fisher, J. A., & Goldin-Meadow, S. (2009). Does linguistic input play the same role in language learning for children with and without early brain injury?.* Developmental Psychology, *45(1), 90.*

found that with preschool-aged children, the density or proportion of sophisticated words used by parents was the strongest predictor of children's vocabulary in kindergarten and second grade, especially when those sophisticated words were used in ways that support understanding. Thus, with infants, helpful lexical input is characterized by more repetition of basic words, by toddlerhood input that is more lexically diverse rather than repetitive is helpful, and by the preschool years, input that is more lexically sophisticated is positively associated with language learning.

We see similar findings with syntactic development, where speech that is more syntactically diverse and complex as children get older predicts children's syntactic skills. Huttenlocher, Vasilyeva, Cymerman, and Levine (2002) found positive associations between the proportion of complex sentences parents used with their 4-years-old at home and the proportion of complex sentences the children used with their parents in those same interactions, but also at school in the classroom. Importantly, they also found positive associations between the proportion of complex sentences that preschool teachers used in their classrooms and growth in syntactic comprehension across the school year of the students in those classes (Huttenlocher et al., 2002), suggesting that input at home and at school can play a role in oral language development. This is an important finding because relations between parent input and child language could just be a byproduct of a genetic effect, but the fact that similar findings exist with teachers and the children in their classes, suggests that it is the environmental exposure that is driving the relation. In a later longitudinal study of parent-child interactions at home in a diverse sample of 47 families in the greater Chicago area, Huttenlocher, Waterfall, Vasilyeva, Vevea, and Hedges (2010) found that between child age 14–46 months parents increased in the diversity of syntactic structures they used in their input, measured as uses of a variety of devices within phrases (adjectives, adverbs, etc.) and as the number of different types of clauses produced (relative clauses, coordinated clause, etc.). Over the period from 26 to 46 months, they found significant positive effects of syntactic diversity on children's syntactic development, suggesting that a greater degree of syntactic diversity in the input over the toddler and preschool years predicts learning during this time. They also found that input mediated SES effects on child language skills, highlighting the important effects of children's language experiences over and above their socioeconomic status (Huttenlocher et al., 2010).

We found parallel findings looking at syntactic complexity rather than diversity and also including a sample of children who had experienced early pre- or perinatal brain injury. We measured the complexity of the parent input as the mean length of the parents' utterances (MLU) in words produced with children during an interaction at child age 30-months and found that parents' MLU positively predicted children's growth in their own mean length of utterance over time (Rowe, Levine, Fisher, & Goldin-Meadow, 2009). This result was even stronger for the group of children who had experienced brain injury compared to the typically developing children. These findings are shown in Fig. 1B. The take-away message here is that even with children who may be delayed in their language skills due to early brain injury, the plasticity is remarkable, and the environmental effects are strong such that in the toddler and preschool years using more diverse and complex utterances in conversations can be helpful. Thus, in most cases there is no need to "pare-down" the input to a lower level.

3.3 A gradual transition from contextualized to decontextualized conversations

In infancy children benefit from talk that is contextualized or focused on the "here and now", such as talking about the ball while playing with the ball. As noted earlier, experimental studies show that infants learn words better when parents label the objects infants are looking at rather than the objects they are not attending to (e.g., Yu & Smith, 2012). The use of gesture in early interactions is another helpful way to discuss the here and now. Deictic gestures such as pointing are often used to direct children's attention to a referent and to reinforce the message conveyed in the spoken input—point at picture of DOG while saying "DOG" (Özçalişkan & Goldin-Meadow, 2005). Studies suggest that when interacting with 1-year-olds, pointing can be helpful for language learning. For example, in one study (Rowe & Goldin-Meadow, 2009) we examined the use of gesture and speech in parent-child interactions when children were 14-months-old in the same socioeconomically diverse sample of families from the greater Chicago area discussed earlier (e.g., Huttenlocher et al., 2010; Rowe et al., 2009). Then we measured the children's vocabulary comprehension skills with the Peabody Picture Vocabulary Test (PPVT III, Dunn & Dunn, 1997) at age 54 months. There were several relevant findings. First, children who gestured to more referents in their environment at 14 months, had higher scores on the PPVT just prior to kindergarten entry than children who gestured less. This result held controlling for how many words the children spoke in the 14-month interaction (which wasn't many).

Furthermore, the significant effect of SES (measured as parent education and family income) on 54-month PPVT scores was significantly partially mediated by children's gesture use at 14 months. This means that some of the socioeconomic disparities in kindergarten vocabulary skill were already evident and explained by variation in early gesture use. Finally, and most importantly for the current discussion, parent use of gesture with children at 14 months significantly predicted child use of gesture at that same time, controlling for SES. These results suggest that by 14 months of age children are already socialized to use more or less gesture during interactions based on their parents' uses of gesture, and this variability in early gesture use predicts later vocabulary size, likely because it elicits parent talk about the child's focus of attention in the here and now (Goldin-Meadow, Goodrich, Sauer, & Iverson, 2007; Rowe & Goldin-Meadow, 2009).

On the other hand, by the late toddler and preschool years when children are more advanced in their cognitive and language skills, they benefit more from talk that is decontextualized rather than contextualized. Decontextualized talk is often defined as talk about the "there and then," for example, talk about the past or the future, or abstract talk such as explanations of how things work in the world (Snow, 1991; Snow, Tabors, & Dickinson, 2001). Just as parents use of gesture with 14-month-olds predicts the children's later vocabulary, by child age 30-months, parents' use of decontextualized talk predicts children's own decontextualized language as well as their later vocabulary, syntax and narrative skills, controlling for parents' use of contextualized talk and SES (Demir et al., 2015).

Why might decontextualized talk be helpful for oral language development? There are several reasons. First, decontextualized utterances, on average, tend to be longer and more syntactically complex than contextualized utterances since more words are often necessary to discuss the "there and then" rather than the "here and now" (e.g., Demir et al., 2015). For example, at the dinner table a parent might use a decontextualized utterance "who was the teacher helper today during school?" and a contextualized utterance "please pass the ketchup". The increased linguistic complexity of decontextualized talk should be helpful because it provides exposure to more diverse and complex structures in the input which we know predicts learning (e.g., Huttenlocher et al., 2010; Rowe et al., 2009). In this way, it also provides children with opportunities to practice using the type of academic language they will encounter in school settings—both in the classroom learning environment and when reading academic texts (Snow, 1991; Snow & Uccelli, 2009). Further, our results show that early use of decontextualized utterances by parents and children in those 30-months interactions, predicts children's 7th grade academic language skills on the Core Academic Language Skills assessment (Uccelli et al., 2015; Uccelli, Demir, Rowe, Levine, & Goldin-Meadow, 2018), controlling for SES, child PPVT at school entry and the amount children talked in the 30-months interaction. Finally, an interesting parallel discussion of this phenomenon has been discussed by O'Madagain and Tomasello (2021) as a transition from discussing a joint focus of attention to objects and events in infancy to a joint focus of attention to mental content (beliefs, desires, opinions) during the preschool years. They theorize that this ability to jointly focus on mental content predicts false belief understanding and reflective reasoning, skills also useful for school success. In sum, a gradual increase over time in the use of decontextualized talk that mirrors the developing zone of proximal development as children move from infancy to preschool age is likely helpful for promoting oral language and literacy development and will help children prepare for the academic language demands of formal schooling.

4. Parenting factors that predict communicative environments and child language development

Young children's early communicative environments vary in the extent to which they contain the helpful features of communication described in the themes presented here. While some of that variation is due to SES (e.g., Hoff, 2003; Huttenlocher et al., 2010), we also see marked variation within SES groups (Huttenlocher et al., 1991; Pan, Rowe, Singer, & Snow, 2005; Weisleder & Fernald, 2013). Furthermore, it is not easy to change a family's SES, thus identifying other factors that predict variation in communicative environments and children's language development is a useful endeavor, especially if it helps to explain the mechanisms underlying SES effects while simultaneously uncovering factors that might be more malleable than SES. Much research has addressed this issue (see Rowe, 2018 for a review), and many factors have been identified. Here I discuss the two that I find most relevant for early intervention and policy targeting young children's communicative environments: Parenting knowledge and parenting stress.

4.1 Parenting knowledge

Parenting knowledge is positively associated with socioeconomic status (Tamis-LeMonda, Shannon, & Spellmann, 2002) and how parents communicate with their young children (e.g., Vernon-Feagans et al., 2008).

For example, Leung and Suskind (2020) found that parents who knew more about infant development when their children were 1-week-old, were more likely to engage in ways that fostered emotional and cognitive growth when the children were 9-months old, and that knowledge of infant development at child age 1-week mediated SES effects on later caregiving behaviors (Leung & Suskind, 2020). Similarly, in the longitudinal Chicago sample mentioned earlier we examined parenting knowledge and parents' verbal intelligence as predictors of parents' communication with children. When the children were 30-months old, we measured the parents' knowledge of child development using the Knowledge of Infant Development Inventory (KIDI, MacPhee, 2002) and the parents' verbal intelligence using the vocabulary subscale of the Weschler Adult Intelligence Scales (WAIS: Weschler, 1981). At this same visit the children were administered the PPVT and the parent and child were videotaped interacting, doing what they would normally do, for 90 min. Using the transcripts of the speech from that video, we made a composite measure of several of the helpful features of input parents' used including their syntactic complexity measured with Mean Length of Utterance (MLU), their vocabulary diversity as well as amount of talk. We found, first of all, that the parent communication composite predicted growth in children's vocabulary in that it was positively associated with the children's PPVT scores 1 year later, controlling for their PPVT scores at 30-months. Second, we found that parent socioeconomic status measured as parent education and family income was positively associated with the parent communication composite such that higher-SES parents used more talk and more diverse and complex talk than lower-SES parents. Importantly, while we found no additional effect of parents' verbal intelligence, we did find that the relationship between SES and parent communication was mediated by parents' knowledge of child development suggesting that a reason there is an association between SES and parent communication is because the higher SES parents, on average, know more about child development which leads them to communicate with their children in more linguistically complex and diverse ways (Rowe, 2008). This mediation effect of parenting knowledge is displayed visually in Fig. 2. Building on this finding, in a larger-scale study with the Early Childhood Longitudinal Study–Birth cohort national dataset, we found that maternal knowledge of child development measured at child age 9-months (using items on the same KIDI measure) showed a direct relationship to child language and pre-literacy skills at age 4, and significantly mediated the relationship between parent education and children's language and literacy outcomes (Rowe, Denmark, Harden, & Stapleton, 2016).

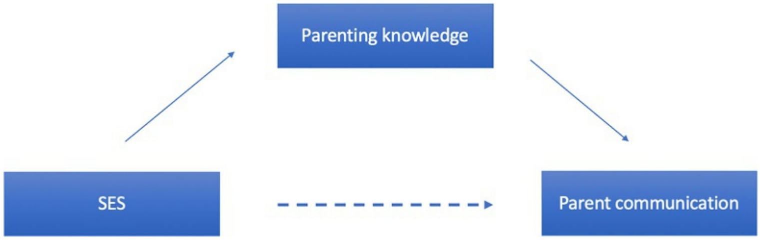


Fig. 2 Schematic drawing of the effect of parenting knowledge and mindsets mediating the relation between SES and parenting communication with children.

One important component of parenting knowledge is the extent to which parents believe their children can improve in their learning and abilities with effort and experiences. Drawing on Dweck's (2006) work on mindsets or theories of intelligence, which posits that individuals hold beliefs that intelligence is either a fixed trait or is malleable and can improve with effort (Dweck, 2006), we examined parents' beliefs about the fixedness of their children's abilities in reading and math (Muenks, Miele, Ramani, Stapleton, & Rowe, 2015). Specifically, we created a scale similar to Dweck's (2006), but more specific to parents' beliefs about their own child's abilities and we looked at associations between parents' fixedness beliefs and the extent to which they reported engaging in different types of activities at home with their child. We found significant associations between parents' fixedness beliefs and their practices such that parents who believed their child's abilities were more "fixed" were more likely to engage in controlling parenting behaviors whereas parents who believed their children's abilities were more malleable were more likely to engage in autonomy supporting behaviors. Our findings were based on parent report (Muenks et al., 2015), but similar results were found in a study that experimentally manipulated parents' mindsets about child ability to be either more fixed or more growth oriented and then examined how parents interacted with their children in a puzzle task. Indeed, the parents who were induced to hold more fixed mindsets engaged in more unconstructive involvement in the task compared to the parents induced to hold growth mindsets (Moorman & Pomerantz, 2010). These studies suggest that mindsets about child ability matter for parenting and are potentially malleable (e.g., Mueller, Rowe, & Zuckerman, 2017).

Additional evidence of the importance of parenting knowledge comes from intervention studies that take the approach of providing parents with knowledge in the form of information or coaching to enhance children's early communicative environments—in essence testing the causal relationships between parenting knowledge, parents' communication with children, and child language outcomes. In general, parent-implemented language interventions of this sort are found to result in significant positive effects, at least in the short term, on parent language behaviors (e.g., Leech, Wei, Harring, & Rowe, 2018; Leung, Hernandez, & Suskind, 2020; Ramírez, Lytle, & Kuhl, 2020; Rowe & Leech, 2019; Suskind et al., 2016) and on child language outcomes (e.g., see Heidlage et al., 2020 for a meta-analysis). In our work we've been taking the approach of targeting the specific features of parent communication highlighted in the themes discussed earlier. As one example, we have a parent training, called *Pointing to Success*, about the usefulness of pointing for infants' developing language skills. We've implemented this training to date in two different relatively small samples and in both cases, there was significant change in parent use of pointing after the training (Rowe & Leech, 2019; Salo et al., under review). In one of the studies, we found effects of the parent training, which took place at child age 10-months, on children's later vocabulary at 18-months, but the effect was moderated by parents' intelligence mindsets such that the intervention had a stronger effect for parents whose mindsets were more "fixed" at baseline than for those who had more growth-oriented mindsets. It may be that the parents with fixed mindsets were just not aware that their communication with their children could make a difference, and thus the intervention had a stronger effect for them once they were informed. Indeed, the intervention itself emphasized the important role of caregiver communication, particularly gesture, in child language learning (e.g., Rowe & Leech, 2019). Similar findings were evident in a larger scale literacy intervention study in Denmark that also took a growth-mindset approach in the intervention materials. Specifically, Andersen and Nielsen (2016) found that providing parents of second graders with children's books and emphasizing the value of supporting their children at home when learning to read, resulted in greater literacy outcomes for the children, especially the children of parents with more fixed mindsets at the start of the study. These studies provide support for interventions that target parent knowledge by highlighting for parents the malleability of child skills and the role that parents can play in promoting those skills. Indeed, parenting knowledge affects parenting practices which affects child outcomes (e.g., Bornstein & Cheah, 2006).

4.2 Parenting stress

Many studies find that stress about parenting itself "parenting stress" (e.g., Deater-Deckard, 2008; Rodgers, 1998) and stress and depression more generally in parents' lives, are negatively related to positive parenting behaviors and to child development. Lower socioeconomic status is associated with higher parenting stress and parental distress. The Family Stress Model (e.g., Conger, Conger, & Martin, 2010) posits a theoretical process by which economic hardships affect child outcomes through the effects on parenting stress and parenting behaviors (e.g., Masarik & Conger, 2017). Many large and small-scale studies confirm this pathway of effects.

For example, Gershoff and colleagues found that parenting stress was significantly negatively associated with positive parenting behaviors in a large national sample of American families, and that stress partially mediated effects of income and hardship on positive parenting behaviors (Gershoff, Aber, Raver, & Lennon, 2007). A separate large-scale study found that maternal depression in particular mediated effects between economic hardship and maternal sensitivity and supportive parenting with 1–3 years-old (Newland, Crnic, Cox, & Mills-Koonce, 2013), and others have found direct effects between parenting stress and positive qualities of parent–child interaction (Farmer & Lee, 2011). Smaller-scale studies have looked more closely at parenting stress and parent–child communication and child language development. One study of a sample of 56 lower-SES families, found a significant relation between parenting stress and preschoolers' expressive and receptive vocabulary (Noel, Peterson, & Jesso, 2008), and suggested the relation might be explained by parent–child communicative interactions. Indeed, our results confirmed this hypothesis. In a longitudinal study of 108 lower-SES families from southern Vermont, we modeled the change over time in parents' talk with their children during play sessions at home at child ages 14, 24 and 36 months. We found that maternal depression was significantly negatively associated with the amount that mothers talked with their children over this time-period (Rowe, Pan, & Ayoub, 2005). Importantly, while much of the research on parent–child interaction and child language development is focused on mothers, there is clear evidence that fathers' speech with children also relates to child language development (Pancsofar & Vernon-Feagans, 2006) and fathers' parenting stress also predicts child language outcomes (e.g., Harewood, Vallotton, & Brophy-Herb, 2017). In sum, evidence suggests that economic hardships associated with lower-socioeconomic status can lead to parenting stress and depression which impact how parents communicate with their children on a daily basis, and ultimately child language development.

Finally, a recent study found that maternal growth mindset played a protective role in the relation between parenting stress and infants' neurodevelopment. More specifically, the study measured maternal stress, maternal mindset, and child neurodevelopment at age 12 months using EEG. Negative associations between maternal stress and child neurodevelopment were found, but only for mothers who held a more fixed mindset, not for mothers with more growth mindsets (Elansary et al., 2021), indicating that having a growth mindset can potentially buffer negative effects of stress. Thus, parenting programs and interventions that promote growth mindset, such as those discussed earlier, are warranted.

5. Implications for social policy and education

Based on the literature, social policies that reduce parenting stress and increase parenting knowledge have the potential to improve early language environments and lead to better outcomes for all children. Here I focus on two reforms that could have a profound positive impact in the US context: Enhancing parental leave plans and teaching parenting and child development in high schools.

5.1 Social policies should enhance parental leave to reduce stress and increase time spent with infants

One issue that adds stress to new parents' lives is the minimal parental leave offered in the United States. New mothers need to balance work obligations with the intense care necessary for a newborn. Evidence suggests that increasing paid family leave will reduce parenting stress, increase and improve parent–child interactions, and lead to improved child health and development.

Some evidence of the positive effects of parental leave policies comes from countries that embrace them. For example, a cross-national study of Canada, the USA, Norway and Germany, found that in each country more educated mothers spent more time with their children. However, in Norway there was no association between fathers' education and time spent with children (and smaller effects were seen in Germany compared to the US). Norway offers the most parental leave to fathers compared to other countries, suggesting that the better parental leave for both parents reduced time constraints on fathers, erasing education effects seen elsewhere (Sayer, Gauthier, & Furstenberg, 2004). Similarly, the Wordbank project (Frank, Braginsky, Yurovsky, & Marchman, 2021) found interesting differences in the magnitude of the effect of parent education on child vocabulary when comparing across 10 countries. It turns out, as an example, that the effect is half the size in Norway compared to the US, suggesting that structural factors such as parental leave can moderate the size of SES effects on child language development (e.g., Rowe & Weisleder, 2020). Norway and the US represent two different ends of the continuum of parental leave policies, as an average American might receive 10–12 weeks of parental leave (often not paid), while parents in Norway receive 12 months paid. While the studies above are mostly correlational, one study from Norway retrospectively estimated the effect of Norway's parental leave reform by examining differences in outcomes for mothers giving birth just before and just after the reform went into place in 1977. Results showed that the increased time parents were able to spend with their child led to a 2% increase in high school graduation rates and 5% increase in wages by age 30. Further, results were strongest for families with lower levels of education (Carneiro, Løken, & Salvanes, 2015).

We also see evidence from within the US. One study based on California's paid family leave plan showed improvements in parental mental health and child health as a result of the plan (Bullinger, 2019) confirming that parental leave can reduce parental stress and depression. Another study using data from the Early Childhood Longitudinal Study–Birth Cohort looked at variation within the US in length of maternity leave. In this sample of 3850 working mothers, the length of leave in weeks ranged from 1 to 52 with a mean of 12.5, and that variation in length of maternity leave was positively associated with the quality of parent-child interactions, controlling for SES (Plotka & Busch-Rossnagel, 2018). In sum, there is growing evidence that enabling parents to spend more time with their newborn children without having to give up employment or lose wages results in reduced parenting stress, and positive parenting and child outcomes.

5.2 Parenting and child development should be taught in high schools

Given the documented relations between parenting knowledge of child development and children's early communicative environments and language development, why not take a preventative approach and teach parenting and child development in secondary schools before most individuals become parents? This would be a cost-effective way to disseminate important knowledge that would help all citizens be more aware of what children are capable of and how to interact with children to promote learning. Further, this would eliminate some need for expensive parenting or multi-generational interventions (St Pierre, Lazer, & Barnes, 1995), as parents would enter into the parenting role holding the knowledge they need to provide optimal learning environments for their children. My colleague Eleanor O'Donnell Weber and I have been pursuing efforts to this end. For example, as a first step, in her dissertation work, O'Donnell Weber (2019) analyzed state standards in the US and determined that in fact approximately half of the U.S. states have standards related to parenting, child development or both, however widespread teaching of these topics is not evident. Thus, this idea is not new, yet the historical and political reasoning behind the standards in different states may differ (e.g., Noddings, 2014).

There is little research on the efficacy of teaching child development in high school, but the studies that do exist are promising in that they suggest students who have taken such courses are better able to judge their own preparedness for parenting (Meyer, Jain, & Canfield-Davis, 2011). This is interesting as some might think teaching high schoolers about parenting and child development would make them want to be parents right away, but in fact it appears to have an opposite effect as it helps them consider the role more realistically. To examine what adolescents know about parenting and child development, O'Donnell Weber (2019) then developed a measure called the Adolescent Parenting Knowledge and Attitudes Survey (APKAS) and administered it online to over 1000 American high school students around the country. The measure consisted of 78 questions on a variety of topics including beliefs about: (1) the importance of active learning for children, (2) the importance of empathic awareness and social-emotional learning, (3) the importance of holding and promoting a growth mindset, (4) supporting oral language development, (5) parenting efficacy, (6) the role parents can play in early learning, as well as questions that tapped into the adolescent's knowledge of typical patterns of child developmental milestones in young American children. She found overall, on the items measuring beliefs, that most of the sample of high schoolers (60%–80%) held beliefs associated with healthy development of children, yet on the survey they tended to "agree" with true statements rather than "strongly agree" indicating room for a shift in their parenting attitudes and beliefs. Regarding parenting knowledge, the results were much less positive, with high school students having very limited knowledge of general developmental milestones (for example at what age a child might say their first word) often indicating that they "don't know." There were 6 questions in this knowledge category and on average the students got fewer than half of them correct. Some interesting further analyses revealed that: Boys were more likely to answer incorrectly than girls or to say "I don't know"; respondents from states that have standards related to parenting and child development scored significantly higher on the knowledge of child development questions than respondents from states without the standards; students who had taken a class on child development, parenting or babysitting scored higher on the knowledge scale, and adolescents who had more experience with child care themselves scored higher on both the beliefs and knowledge measures than adolescents with less child-care experience (O'Donnell Weber, 2019). Taken together these results suggest that American adolescents do not have much knowledge of child development, on average, and those who have taken a course, know more.

In sum, we see great potential value in teaching high school students about how to foster development and learning in young children. Most Americans will become parents in their lifetime (Newport & Wilke, 2013) and American parents are most likely to turn to their own family members when in need of parenting information (e.g., McCatharn, Herbert, Wei, & Rowe, 2021; Rowe et al., 2016) However, more educated parents report more often going to professionals (e.g., pediatricians) for parenting information, and parents who report going to professionals for parenting information are more likely to have more parenting knowledge (Rowe et al., 2016). In the study mentioned above the adolescents were also asked where they learned about parenting and the most common response was "from watching their own parents" with the second most common response "from movies or television" (O'Donnell Weber, 2019). This suggests that in addition to directly providing students with knowledge of child development in schools, the media industry could embrace the goal of providing examples of positive parenting in movies, television and social media.

Given the incredible plasticity during early childhood and the importance of children's environments during this time, preparing parents ahead of time for the parenting role can help promote early learning and save costs associated with parenting interventions. Indeed, most existing parenting intervention programs are costly, and limited in scope and population (e.g., St Pierre et al., 1995). Therefore, a greater way to scale information about child development is through the school system. Whether this content should be taught in its own course, which may be hard to require given curricular pressures, or across disciplines through incorporating positive examples of parenting through relevant literature and content (e.g., Noddings, 2014) is an open and interesting question.

6. Conclusions

From birth, children's early communicative environments can provide rich opportunities to build oral language skills that serve as a foundation for learning to read and school success. Here, I presented several themes that highlight the features of these early communicative environments that can be particularly helpful for promoting language learning. The first is a focus on engaging the child in back-and-forth conversations. The second theme highlights how these conversations can become increasingly more sophisticated linguistically through including more diverse and complex words and structures as children get older. The third theme underlines how the content of the conversations should also change over early development from a focus on more contextualized or "here and now" topics to more abstract or decontextualized topics. Variation in these qualities of children's early communicative environments predict language development, over and above family SES, and can be explained at least in part by parents' knowledge of child development and parenting stress. Targeting parenting knowledge and parenting stress through policy could improve early communicative environments and child language and literacy development. Two reforms that could be particularly effective as discussed here include improving and extending parental leave policies in the US and teaching parenting and child development in American high schools.

References

Andersen, S. C., & Nielsen, H. S. (2016). Reading intervention with a growth mindset approach improves children's skills. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(43), 12111–12113.

Bornstein, M. H., & Cheah, C. S. L. (2006). The place of 'culture and parenting' in the ecological contextual perspective on developmental science. In K. H. Rubin, & O. B. Chung (Eds.), *Parenting beliefs, behaviors, and parent–child relations* (pp. 3–33). New York: Psychology Press.

Brito, N. H., Fifer, W. P., Myers, M. M., Elliott, A. J., & Noble, K. G. (2016). Associations among family socioeconomic status, EEG power at birth, and cognitive skills during infancy. *Developmental Cognitive Neuroscience*, *19*, 144–151.

Bullinger, L. R. (2019). The effect of paid family leave on infant and parental health in the United States. *Journal of Health Economics*, *66*, 101–116.

Carneiro, P., Løken, K. V., & Salvanes, K. G. (2015). A flying start? Maternity leave benefits and long-run outcomes of children. *Journal of Political Economy*, *123*(2), 365–412.

Conger, R. D., Conger, K. J., & Martin, M. J. (2010). Socioeconomic status, family processes, and individual development. *Journal of Marriage and Family*, *72*(3), 685–704.

Deater-Deckard, K. (2008). *Parenting stress*. Yale University Press.

Demir, Ö. E., Rowe, M. L., Heller, G., Goldin-Meadow, S., & Levine, S. C. (2015). Vocabulary, syntax, and narrative development in typically developing children and children with early unilateral brain injury: Early parental talk about the "there-and-then" matters. *Developmental Psychology*, *51*(2), 161.

Dickinson, D. K., Golinkoff, R. M., & Hirsh-Pasek, K. (2010). Speaking out for language: Why language is central to reading development. *Educational Researcher*, *39*(4), 305–310.

Donnellan, E., Bannard, C., McGillion, M. L., Slocombe, K. E., & Matthews, D. (2020). Infants' intentionally communicative vocalizations elicit responses from caregivers and are the best predictors of the transition to language: A longitudinal investigation of infants' vocalizations, gestures and word production. *Developmental Science*, *23*(1), e12843.

Duff, F. J., Reen, G., Plunkett, K., & Nation, K. (2015). Do infant vocabulary skills predict school-age language and literacy outcomes? *Journal of Child Psychology and Psychiatry*, *56*(8), 848–856.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody picture vocabulary test* (3rd ed.). Circle Pines.

Durham, R. E., Farkas, G., Hammer, C. S., Tomblin, J. B., & Catts, H. W. (2007). Kindergarten oral language skill: A key variable in the intergenerational transmission of socioeconomic status. *Research in Social Stratification and Mobility*, *25*(4), 294–305.

Dweck, C. S. (2006). *Mindset: The new psychology of success*. Random House.

Elansary, M., Pierce, L. J., Wei, W. S., McCoy, D. C., Zuckerman, B., & Nelson, C. A. (2021). Maternal stress and early neurodevelopment: Exploring the protective role of maternal growth mindset. *Journal of Developmental and Behavioral Pediatrics: JDBP*.

Farmer, A. Y., & Lee, S. K. (2011). The effects of parenting stress, perceived mastery, and maternal depression on parent–child interaction. *Journal of Social Service Research*, *37*(5), 516–525.

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, *42*(1), 98–116.

Foorman, B. R., Herrera, S., Petscher, Y., Mitchell, A., & Truckenmiller, A. (2015). The structure of oral language and reading and their relation to comprehension in Kindergarten through Grade 2. *Reading and Writing*, *28*(5), 655–681.

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2021). *Variability and consistency in early language learning: The Wordbank project*. MIT Press.

Garcia, E., & Weiss, E. (2017). *Education inequalities at the school starting gate: Gaps, trends, and strategies to address them*. Economic Policy Institute.

Gershoff, E. T., Aber, J. L., Raver, C. C., & Lennon, M. C. (2007). Income is not enough: Incorporating material hardship into models of income associations with parenting and child development. *Child Development*, *78*, 70–95.

Goldin-Meadow, S., Goodrich, W., Sauer, E., & Iverson, J. (2007). Young children use their hands to tell their mothers what to say. *Developmental Science*, *10*(6), 778–785.

Golinkoff, R. M., Hoff, E., Rowe, M. L., Tamis-LeMonda, C. S., & Hirsh-Pasek, K. (2019). Language matters: Denying the existence of the 30-million-word gap has serious consequences. *Child Development*, *90*(3), 985–992.

Harewood, T., Vallotton, C. D., & Brophy-Herb, H. (2017). More than just the breadwinner: The effects of fathers' parenting stress on children's language and cognitive development. *Infant and Child Development*, *26*(2), e1984.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Paul H Brookes Publishing.

Heckman, J. J. (2011). The economics of inequality: The value of early childhood education. *American Educator*, *35*(1), 31.

Heidlage, J. K., Cunningham, J. E., Kaiser, A. P., Trivette, C. M., Barton, E. E., Frey, J. R., et al. (2020). The effects of parent-implemented language interventions on child linguistic outcomes: A meta-analysis. *Early Childhood Research Quarterly*, *50*, 6–23.

Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., et al. (2015). The contribution of early communication quality to low-income children's language success. *Psychological Science*, *26*(7), 1071–1083.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, *74*(5), 1368–1378.

Huttenlocher, P. R. (2009). *Neural plasticity*. Harvard University Press.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*(2), 236.

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology*, *45*(3), 337–374.

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, *61*(4), 343–365.

Kuhl, P. K. (2007). Is speech learning 'gated' by the social brain? *Developmental Science*, *10*(1), 110–120.

Leech, K. A., Wei, R., Harring, J., & Rowe, M. L. (2018). A brief parent-focused intervention to improve preschoolers' conversational skills and school readiness. *Developmental Psychology*, *54*, 15–28.

Leung, C. Y., Hernandez, M. W., & Suskind, D. L. (2020). Enriching home language environment among families from low-SES backgrounds: A randomized controlled trial of a home visiting curriculum. *Early Childhood Research Quarterly*, *50*, 24–35.

Leung, C. Y., & Suskind, D. L. (2020). What parents know matters: Parental knowledge at birth predicts caregiving behaviors at 9 months. *The Journal of Pediatrics*, *221*, 72–80.

MacPhee, D. (2002). *Knowledge of infant development inventory: Manual*. Colorado: Colorado State University.

Maguire, M. J., Schneider, J. M., Middleton, A. E., Ralph, Y., Lopez, M., Ackerman, R. A., et al. (2018). Vocabulary knowledge mediates the link between socioeconomic status and word learning in grade school. *Journal of Experimental Child Psychology*, *166*, 679–695.

Masarik, A. S., & Conger, R. D. (2017). Stress and child development: A review of the family stress model. *Current Opinion in Psychology*, *13*, 85–90.

McCatharn, J. M., Herbert, K. K., Wei, R., & Rowe, M. L. (2021). Circles of support: Exploring the 'Where' and 'Why' of Parents' of infants information seeking behaviors. *Journal of Child and Family Studies*, 1–14.

Meyer, B. W., Jain, S., & Canfield-Davis, K. (2011). The effect of parenthood education on self-efficacy and parent effectiveness in an alternative high school student population. *The Professional Counselor*, *1*(1), 29–40.

Moorman, E. A., & Pomerantz, E. M. (2010). Ability mindsets influence the quality of mothers' involvement in children's learning: An experimental investigation. *Developmental Psychology*, *46*(5), 1354.

Mueller, C., Rowe, M. L., & Zuckerman, B. (2017). Mindset matters for parents and adolescents. *JAMA Pediatrics*, *171*(5), 415–416.

Muenks, K., Miele, D. B., Ramani, G. B., Stapleton, L. M., & Rowe, M. L. (2015). Parental beliefs about the fixedness of ability. *Journal of Applied Developmental Psychology*, *41*, 78–89.

Newland, R. P., Crnic, K. A., Cox, M. J., & Mills-Koonce, W. R. (2013). The family model stress and maternal psychological symptoms: Mediated pathways from economic hardship to parenting. *Journal of Family Psychology*, *27*(1), 96.

Newman, R. S., Rowe, M. L., & Ratner, N. B. (2016). Input and uptake at 7 months predicts toddler vocabulary: The role of child-directed speech and infant processing skills in language development. *Journal of Child Language*, *43*(5), 1158–1173.

Newport, F., & Wilke, J. (2013). *Desire for children still norm in U.S.* Retrieved September 24, 2018, from https://news.gallup.com/poll/164618/desire-children-norm.aspx.

NICHD Early Child Care Research Network. (2005). Pathways to reading: The role of oral language in the transition to reading. *Developmental Psychology*, *41*(2), 428–442.
Noble, K. G., Houston, S. M., Kan, E., & Sowell, E. R. (2012). Neural correlates of socioeconomic status in the developing human brain. *Developmental Science*, *15*(4), 516–527.
Noddings, N. (2014). Can we teach parenting in our schools? *Theory and Research in Education*, *12*(2), 221–228.
Noel, M., Peterson, C., & Jesso, B. (2008). The relationship of parenting stress and child temperament to language development among economically disadvantaged preschoolers. *Journal of Child Language*, *35*(4), 823–843.
O'Donnell Weber, E. B. (2019). *American high school Students' knowledge and beliefs about parenting and child development*. (doctoral dissertation).
O'Madagain, C., & Tomasello, M. (2021). Joint attention to mental content and the social origin of reasoning. *Synthese*, *198*(5), 4057–4078.
Özçalişkan, Ş., & Goldin-Meadow, S. (2005). Do parents lead their children by the hand? *Journal of Child Language*, *32*(3), 481–505.
Pan, B. A., Rowe, M. L., Singer, J. D., & Snow, C. E. (2005). Maternal correlates of growth in toddler vocabulary production in low-income families. *Child Development*, *76*, 763–782.
Pancsofar, N., & Vernon-Feagans, L. (2006). Mother and father language input to young children: Contributions to later language development. *Journal of Applied Developmental Psychology*, *27*(6), 571–587.
Plotka, R., & Busch-Rossnagel, N. A. (2018). The role of length of maternity leave in supporting mother–child interactions and attachment security among American mothers and their infants. *International Journal of Child Care and Education Policy*, *12*(1), 1–18.
Ramírez, N. F., Lytle, S. R., & Kuhl, P. K. (2020). Parent coaching increases conversational turns and advances infant language development. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(7), 3484–3491.
Reardon, S. F. (2013). The widening income achievement gap. *Educational Leadership*, *70*(8), 10–16.
Rodgers, A. Y. (1998). Multiple sources of stress and parenting behavior. *Children and Youth Services Review*, *20*(6), 525–546.
Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., et al. (2018). Beyond the 30-million-word gap: Children's conversational exposure is associated with language-related brain function. *Psychological Science*, *29*(5), 700–710.
Rowe, M. L. (2008). Child-directed speech: Relation to socioeconomic status, knowledge of child development, and child vocabulary skill. *Journal of Child Language*, *35*, 185–205.
Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development*, *83*(5), 1762–1774.
Rowe, M. L. (2018). Understanding socioeconomic differences in parents' speech to children. *Child Development Perspectives*, *12*(2), 122–127.
Rowe, M. L., Denmark, N., Harden, B. J., & Stapleton, L. M. (2016). The role of parent education and parenting knowledge in children's language and literacy skills among White, Black, and Latino families. *Infant and Child Development*, *25*(2), 198–220.
Rowe, M. L., & Goldin-Meadow, S. (2009). Differences in early gesture explain SES disparities in child vocabulary size at school entry. *Science*, *323*(5916), 951–953.
Rowe, M. L., & Leech, K. A. (2019). A parent intervention with a growth mindset approach improves children's early gesture and vocabulary development. *Developmental Science*, *22*(4).
Rowe, M. L., Levine, S. C., Fisher, J. A., & Goldin-Meadow, S. (2009). Does linguistic input play the same role in language learning for children with and without early brain injury? *Developmental Psychology*, *45*(1), 90.

Rowe, M. L., Pan, B. A., & Ayoub, C. (2005). Predictors of variation in maternal talk to children: A longitudinal study of low-income families. *Parenting: Science and Practice*, *5*(3), 259–283.

Rowe, M. L., & Weisleder, A. (2020). Language development in context. *Annual Review of Developmental Psychology*, *2*, 201–223.

Salo, V.C., Debnath, R., Rowe, M.L., Fox, N.A. (under review). Gesture experience facilitates infant vocabulary growth through enhancement of sensorimotor brain activity.

Sayer, L. C., Gauthier, A. H., & Furstenberg, F. F., Jr. (2004). Educational differences in parents' time with children: Cross-national variations. *Journal of Marriage and Family*, *66*(5), 1152–1169.

Schwab, J. F., Rowe, M. L., Cabrera, N., & Lew-Williams, C. (2018). Fathers' repetition of words is coupled with children's vocabularies. *Journal of Experimental Child Psychology*, *166*, 437–450.

Shneidman, L. A., & Goldin-Meadow, S. (2012). Language input and acquisition in a Mayan village: How important is directed speech? *Developmental Science*, *15*(5), 659–673.

Shneidman, L., & Woodward, A. L. (2016). Are child-directed interactions the cradle of social learning? *Psychological Bulletin*, *142*(1), 1.

Snow, C. E. (1991). The theoretical basis for relationships between language and literacy in development. *Journal of Research in Childhood Education*, *6*(1), 5–10.

Snow, C. E., Tabors, P. O., & Dickinson, D. K. (2001). Language development in the preschool years. In *Beginning literacy with language: Young children learning at home and school* (pp. 1–25).

Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. *The Cambridge handbook of literacy*, *112*, 133.

St Pierre, R. G., Lazer, J. I., & Barnes, H. V. (1995). Two-generation programs: Design, cost, and short-term effectiveness. *Future of Children*, *5*(3), 76–93.

Suskind, D. L., Leffel, K. R., Graf, E., Hernandez, M. W., Gunderson, E. A., Sapolich, S. G., et al. (2016). A parent-directed language intervention for children of low socioeconomic status: A randomized controlled pilot study. *Journal of Child Language*, *43*(2), 366–406.

Tamis-LeMonda, C. S., Shannon, J., & Spellmann, M. (2002). Low-income adolescent mothers' knowledge about domains of child development. *Infant Mental Health Journal*, *23*(1–2), 88–103.

Tomalski, P., Moore, D. G., Ribeiro, H., Axelsson, E. L., Murphy, E., Karmiloff-Smith, A., et al. (2013). Socioeconomic status and functional brain development–associations in early infancy. *Developmental Science*, *16*(5), 676–687.

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development*, *57*(6), 1454–1463.

Uccelli, P., Barr, C. D., Dobbs, C. L., Galloway, E. P., Meneses, A., & Sánchez, E. (2015). Core academic language skills: An expanded operational construct and a novel instrument to chart school-relevant language proficiency in preadolescent and adolescent learners. *Applied PsychoLinguistics*, *36*(5), 1077–1109.

Uccelli, P., Demir, E., Rowe, M. L., Levine, S., & Goldin-Meadow, S. (2018). Children's early decontextualized talk predicts academic language proficiency in mid-adolescence. *Child Development*.

Vernon-Feagans, L., Pancsofar, N., Willoughby, M., Odom, E., Quade, A., Cox, M., et al. (2008). Predictors of maternal language to infants during a picture book task in the home: Family SES, child characteristics and the parenting environment. *Journal of Applied Developmental Psychology*, *29*(3), 213–226.

Vygotsky, L. S. (1978). *Mind in society: Development of higher psychological processes*. Harvard University Press.

Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, *24*(11), 2143–2152.

Weizman, Z. O., & Snow, C. E. (2001). Lexical output as related to children's vocabulary acquisition: Effects of sophisticated exposure and support for meaning. *Developmental Psychology*, *37*(2), 265.
Weschler, D. (1981). *Weschler adult intelligence scale-revised*. New York: The Psychological Corporation.
Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, *125*(2), 244–262.

This page intentionally left blank

CHAPTER FIVE

Kindness towards all: Prosocial behaviors to address U.S. Latinx youth social inequities

Gustavo Carlo[a,*], George P. Knight[b], and Alexandra N. Davis[c]
[a]School of Education, University of California, Irvine, CA, United States
[b]Arizona State University, Tempe, AZ, United States
[c]University of New Mexico, Albuquerque, NM, United States
*Corresponding author: e-mail address: gcarlo@uci.edu

Contents

Abstract

Latinx youth in the United States face structural barriers that contribute to inequities across multiple domains (e.g., education, juvenile justice, healthcare systems), as racial biases permeate social institutions. The systemic oppression resulting from racism can be seen in disparities across many indicators of health, including physical health, education, socioeconomic conditions, and the overrepresentation of ethnic and racial minority individuals, including Latinx individuals, incarcerated and exposed to violence. We present an approach to combat social inequities and injustices by promoting and fostering prosocial behaviors (i.e., actions that benefit others) between majority and minority members of our society. Existing theories and research on the factors that can promote such behaviors across youth from different backgrounds is summarized

Advances in Child Development and Behavior, Volume 63
ISSN 0065-2407
https://doi.org/10.1016/bs.acdb.2022.04.002

though we highlight work in U.S. Latinx youth. Factors that enhance and undermine prosocial behaviors towards diverse others is also summarized. Finally, some recommendations for intervention and policy efforts are briefly presented.

1. Social injustices and inequities in Latinx youth populations

1.1 Structural and systemic challenges

Latinx youth in the United States face structural barriers that contribute to inequities across multiple domains (e.g., education, juvenile justice, healthcare systems), as racial biases permeate social institutions (see Espinola, Zhen-Duan, Suarez-Cano, Mowry-Mora, & Schultz, 2019). The systemic oppression resulting from racism can be seen in disparities across many indicators of health, including physical health, education, socioeconomic conditions, and the overrepresentation of ethnic and racial minority individuals, including Latinx individuals, incarcerated and exposed to violence (see Espinola et al., 2019). U.S. Latinx families often face institutions and societal policies that limit access to opportunities, such as economic opportunities and high quality education (see Espinola et al., 2019). For example, U.S. Latinx students experience higher "push-out" rates (students are often pushed out of the education system as opposed to simply dropping out; Doll, Eslami, & Walters, 2013) than their White, European American peers and are thus at risk for living in poverty because of restricted educational opportunities (see Espinola et al., 2019). Structural inequalities are pervasive, resulting in experiences of both chronic and acute stress that can impact family processes and youth adjustment (Flores et al., 2008).

Many U.S. Latinx families face chronic stress that results from living in poverty, as Latinx families are disproportionally at risk for living in poverty with limited access to financial and community resources (Berlan & Harwood, 2018). U.S. Latinx families also experience acute stressors associated with navigating U.S. culture and societal systems (e.g., immigration; Berry, 2017). Additionally, the COVID-19 pandemic exacerbated many of the existing stressors for Latinx families because of the financial strain caused by the economic downturn as well as increases in prejudice toward ethnic/racial minority and low-SES populations during this historical event (FBI, 2020). Because of the systemic barriers and stressors faced by

U.S. Latinx families, it is important to understand positive development factors that can promote health and social well-being under conditions of environmental risk.

1.2 Prosocial behaviors as a mechanism of social justice

To address the many social injustice and inequity challenges requires a multilevel approach aimed at changing structural systems, laws and policies, and communities at the broadest level but also implementing programs and policies that foster and promote positive social interactions and relationships among diverse groups. For decades, social and behavioral scholars have studied the development and correlates of prosocial behaviors. Prosocial behaviors are defined as actions intended to benefit others (Carlo & Randall, 2002; see Eisenberg, Fabes, & Spinard, 2006). These actions are quite varied and include sharing, donating, volunteering, comforting others, helping, and defending or rescuing persons in danger. As can be inferred, some of these behaviors are purposeful and structured (e.g., volunteering for a charity organization) and might require planning whereas other behaviors might be impulsive (e.g., rescuing a person in immediate danger) and conducted without much planning. In addition, prosocial actions can be motivated by selfish or instrumental concerns but other prosocial actions might be personally costly (psychologically or physically) and motivated primarily by the concern for others. For example, a person might choose to help someone else as a means to gain someone's trust for practical reasons or to gain social approval or status. In contrast, some persons might help someone in need even at their own personal physical health risk or even if doing so is financially costly. This latter form of prosocial behaviors is referred to as altruistic behaviors. Although enactment of altruistic behaviors can sometimes result in a self-benefit (e.g., improve one's mood), the primary intention is sufficiently powerful to override concerns with one's personal cost of helping. Although there are varied motives and circumstances that move distinct forms of prosocial behaviors, we assert that prosocial behaviors towards diverse others are fundamental actions that mitigate social inequities (see Davis, Carlo, & Maiya, 2021).

Importantly, however, there is ample evidence of individual differences in prosocial actions, and researchers have identified several core personal traits linked to these actions (Carlo, 2014). Individual differences in prosocial behaviors help us to understand the wide variability that exists in persons' willingness to help diverse others even when social inequities and injustices are observed. For example, two strongly related characteristics: empathy and sympathy (Eisenberg, 1986; Staub, 2005) have been associated with multiple forms of prosocial behaviors. Scholars define empathy as feeling (positive or negative valence) the same as another and sympathy as feelings of sorrow, sadness, or concern for another. According to researchers, persons who are moved strongly to empathize or to sympathize with another who is in distress, are more likely to engage in prosocial behaviors behaviors (see Batson, 1998, for a review). Such persons are often primarily motivated to help others in need in order to reduce the needy others' distress.

A second set of characteristics that is closely tied to empathy, sympathy, and prosocial behaviors is strongly internalized or endorsed moral principles and values (Eisenberg, 1986). Principles that place high regard for reducing suffering in others, treating others humanely, caring for others, and belief in treating others equitably with respect, fairness, and justice, are all examples of values that can induce prosocial actions. Some values associated with prosocial actions might be more strongly endorsed in specific cultural groups such as familism (i.e., duty or obligation to, support to and from, and affinity to, the family unit; Knight & Carlo, 2012). Finally, these scholars assert that some altruistic and other prosocial behaviors are moved by both set of traits, empathy/sympathy and internalized principles. Indeed, Eisenberg and Fabes (1991) noted that empathy and sympathy can sometimes induce moral principles and values and vice versa—sometimes, moral principles and values can induce empathy and sympathy (see also Hoffman, 2000).

Given that minoritized groups, including U.S. Latinx youth, are the primary targets of social injustices and inequities and that such groups are often economically, politically, and educationally deprived as a result of a legacy of social and systemic oppression, we assert the critical need to foster and promote prosocial behaviors between majority and minority groups in order to facilitate harmonious intergroup relations that can break down social barriers. Multiple forms of prosocial behaviors have the potential to facilitate intergroup cooperation and positive development for marginalized youth (Taylor, 2020), and therefore, considering prosocial behaviors that occur in diverse situations and with different motivations is important. The present essay presents an overview of scholarship on prosocial development with a focus on the relevance of this work as an avenue towards addressing social injustices and inequities that affect minority youth.

1.3 Prosocial behaviors as a marker of social wellbeing and health

The research on prosocial behaviors in the past decade has revealed the fact that prosocial behaviors are much more than a marker of morality. There is accumulating evidence that prosocial behaviors are also linked to important markers of health including psychological adjustment (e.g., depression, anxiety, self esteem; Carlo, 2014; Davis et al., 2016), physical health (e.g., c-reactive proteins, Schreier, Schonert-Reichl, & Chen, 2013), health behaviors (e.g., exercise, sleep; Spitzer & Hollmann, 2013), academic performance (Carlo, White, Streit, Knight, & Zeiders, 2018), and externalizing behaviors (e.g., illegal substance use, delinquency, aggression, bullying; Carlo et al., 2014; Davis, Carlo, Hardy, Olthuis, & Zamboanga, 2017; Walters, 2020). Moreover, a number of sociocognitive and socioemotional traits, such as perspective taking, moral reasoning, empathy and sympathy, guilt, shame, and interpersonal trust, are related to prosocial behaviors (see Carlo, 2014). Finally, youth who exhibit high levels of prosocial behaviors also exhibit positive interpersonal relationships with parents and peers, and are less exposed to violent media (Carlo, McGinley, Hayes, & Martinez, 2012; see Carlo, 2014). As a whole, then, this body of work shows that prosocial behaviors are a marker of social and behavioral well-being and health.

The existing work thus suggests that prosocial behaviors can enhance personal and social health and well-being. This is important because the challenges of addressing social injustice and inequities require attention to the personal health and social well-being of minority group members. Enhancing minority group members' personal health and social well-being empowers such persons to successfully integrate into their communities. Moreover, the negative link between prosocial behaviors and antisocial behaviors provides an avenue to foster intergroup harmony and cooperation and to simultaneously mitigate selfishly motivated and harmful actions toward outgroup members. This approach, in general, presents a holistic, strengths-based approach to redress social inequities that might affect often-marginalized and vulnerable minority groups. Thus, the development of a prosocial orientation may impact social inequities and disparities over time. Furthermore, the socialization of such prosocial behaviors and orientation may be an essential pathway for addressing social inequities and disparities.

2. Traditional approaches to the study of prosocial development

2.1 Cognitive developmental theories and research

Cognitive developmental theorists (e.g., Piaget, Kohlberg), who highlighted the role of sociocognitive skills such as perspective taking and moral reasoning, guided early developmental studies of prosocial development. Perspective taking refers to understanding the thoughts, feelings, and social circumstances of others and is deemed to develop with age across childhood and adolescence (see Carlo, 2014). Moral reasoning, on the other hand, is how children and adolescents resolve dilemma scenarios where one's needs are in conflict with another's (see Carlo, 2014). In prohibitive dilemma scenarios, the protagonist must decide between violating a law or formal social rule and meeting one's own needs or desires, and sometimes such dilemmas can have serious life-threatening consequences (see Carlo, 2014). In prosocial dilemma scenarios, the tension is between helping someone who is in distress or need and meeting one's own needs or desires (see Carlo, 2014). Importantly, however, in prosocial moral reasoning scenarios, the situation is serious but not life threatening. This element creates dilemmas that are personal decisions and not subject to influence from formal laws or rules. In general, substantive evidence reliably demonstrates that both perspective taking and moral reasoning (especially prosocial moral reasoning) are positively related to prosocial behaviors, including altruistic behaviors (Carlo et al., 2014; Carlo, Knight, McGinley, Zamboanga, & Jarvis, 2010; Eisenberg, Eggum-Wilkens, & Spinrad, 2015).

2.2 Traditional socialization theories and research

A distinct line of research has focused on the roles of socialization agents (e.g., parents, peers, media) and socioemotive traits such as empathy and sympathy (Eisenberg, 1986; Hoffman, 2000; Staub, 1978). Although there is increasing work on the study of various socialization agents (see Carlo, 2014; Eisenberg et al., 2015), the bulk of the research has focused on the influence of parents. Moral socialization scholars have identified several parental practices that are associated with prosocial and moral development. For example, harsh and authoritarian parenting practices, that might include the use of corporal punishment and power assertion, has been linked to low levels of prosocial behaviors (Carlo, Knight, McGinley, & Hayes, 2011). In contrast, supportive and authoritative parenting styles, that often includes inductive (child-centered and the use of explanations) disciplining practices, is related to high levels of prosocial behaviors (Krevans & Gibbs, 1996). Other research has examined the role of helicopter parenting (i.e., McGinley, 2018), proactive parenting (Padilla-Walker, Fraser, & Harper, 2012), prosocial parenting practices (e.g., use of social and material rewards, moral conversations) and their links to prosocial behaviors (Davis & Carlo, 2018). Moreover, these scholars have focused on the central role of empathy and sympathy (see Carlo, 2014) as important predictors of prosocial behaviors, and there is ample supportive evidence for this assertion (Eisenberg et al., 2015).

As the body of work that validated the assertions of both cognitive developmental and moral socialization theorists accumulated, some scholars posited integrative approaches such as social cognitive theories (Bandura, 1986; Eisenberg, 1986; Staub, 1978). These latter theories postulated socialization mechanisms that foster sociocognitive and socioemotive traits that were subsequently linked to prosocial and moral development. Much of the existing research on these traditional socialization approaches, for example, provides general support that parenting predicts children's development of empathy, sympathy, perspective taking, and moral reasoning (see Carlo, 2014; Davis & Carlo, 2018), which in turn, predict prosocial behaviors (Davis & Carlo, 2018; Gülseven & Carlo, 2021).

2.3 Cultural socialization theories and research

Theoretical approaches that acknowledged sociocognitive, socioemotive and social influences were an important advance in prosocial development work. However, given the accumulating evidence of culture group differences in prosocial and altruistic behaviors (see de Guzman, Do, & Kok, 2014; Kumru, Carlo, Mestre, & Samper, 2012), there was a need to also incorporate culture-related mechanisms to account for such differences.

Cultural developmental scholars had noted the impact of culture-related practices, beliefs, and physical settings that were posited to influence child development (Super & Harkness, 1997; Whiting & Whiting, 1975). García-Coll et al. (1996) had also presented an integrative model of ethnic minority youth that outlined systemic and social oppressive forces and culture-related mechanisms that impacted ethnic and racial minority children's development (García-Coll et al., 1996). Importantly, within the field of prosocial development, Knight and his associates (Knight, Bernal, Garza, & Cota, 1993; Knight, Carlo, Mahrer, & Davis, 2016) had

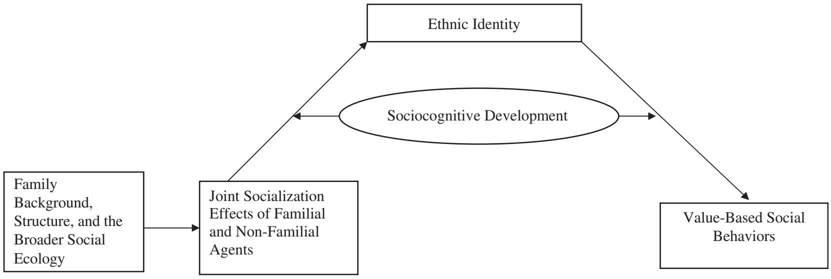


Fig. 1 A cultural socialization model of cooperative and competitive behaviors. *Adapted from Knight, G. P., Bernal. M. E., Garza, C. A., & Cota, M. K. (1993). A social cognitive model of ethnic identity and ethnically-based behaviors. In M. E. Bernal, G. P. Knight (Eds.),* Ethnic identity: Formation and transmission among Hispanics and other minorities *(pp. 213–234). Albany, NY: State University of New York Press.*

demonstrated the central role of ethnic socialization practices, ethnic identity (as a mediator), and sociocognitive skills (as a moderator) in accounting for U.S. Mexican children's cooperative behaviors (see Fig. 1).

Extending these and other (e.g., Laosa & Henderson, 1991) prior theories and models, Carlo and his colleagues (Carlo & Conejo, 2019; Raffaelli, Carlo, Carranza, & Gonzales-Kruger, 2005) proposed an *Ecocultural Model of U.S. Latinx Youth Development* (see Fig. 2) that delineated distal (e.g., school and receiving community context characteristics, major life events, history and origin context of immigration, family and non-familial socialization agents) and specific proximal (e.g., ethnic identity, child discrimination experiences, stress appraisals, perspective taking, moral reasoning, cultural values, empathy, self regulation) influences that impact subsequent U.S. Latinx youth adjustment (including prosocial behaviors).

The approach is founded on the notion that there are culture-related assets and risks that affect U.S. Latinx youth development. Indeed, identification of culture-related mechanisms are necessary to unpack the influence of culture on youth development. Moreover, the approach assumes that there are individual differences in how youth process exposure to the various distal and proximal influences. For example, several traditional Latinx ethnocentric beliefs and concepts, such as familism, *respeto* (i.e., respect), *bien educado* (i.e., good manners and moral character), ethnic identity, and humility, are deemed protective and enhancing factors (Bridges et al., 2012; Knight et al., 2010). In fact, these culture-associated values inherently orient Latinx youth towards prosociality because they guide the youth in being considerate of others and the broader group, and less about themselves.

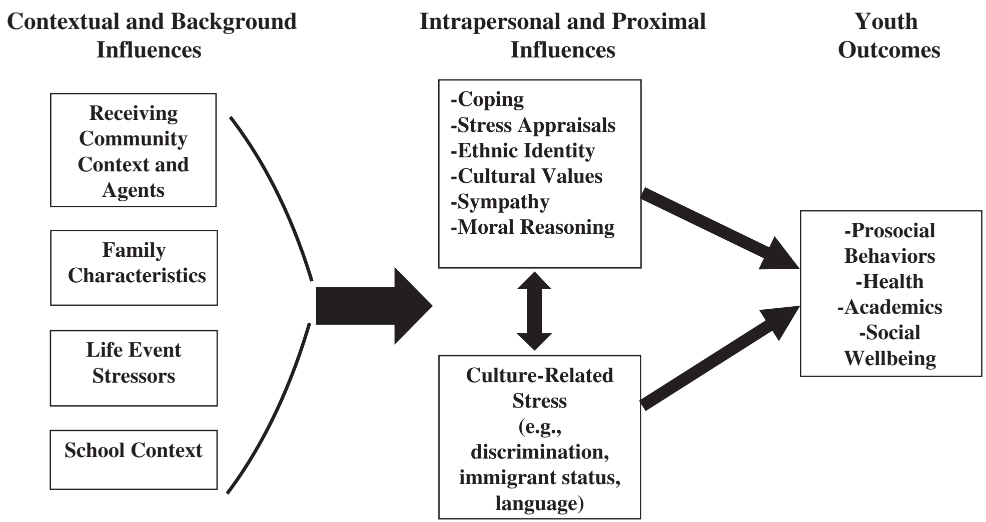


Fig. 2 An ecocultural model of U.S. Latinx Youth Development. *Adapted from Carlo, G., & Conejo, L. D. (2019). Traditional and culture-specific parenting of prosociality in U.S. Latinx youth. In D. Laible, L. Padilla-Walker, & G. Carlo (Eds.),* Oxford handbook of parenting and moral development. *New York: Oxford University Press.*

In addition, because these concepts are highly valued in most Latinx families, there are familial culture-based socialization practices that are designed to transmit these values and notions to their youth.

Of the various culture-related resiliency factors, perhaps the two most studied are familism and ethnic identity. Several studies show positive relations between familism and care-based, other-oriented prosocial behaviors in U.S. Latinx youth (Knight, Carlo, Basilio, & Jacobson, 2015; Streit, Carlo, Killoren, & Alfaro, 2018). Knight, Mazza, and Carlo (2018) show empirical relations of developmental trajectories of familism (presumably being developed through ethnic socialization experiences) to prosocial behaviors. Interestingly, familism has been linked to prosocial behaviors via its effects on perspective taking and prosocial moral reasoning (Knight et al., 2015). Such findings suggest that cultural values can sometimes enhance prosocial and moral tendencies associated with prosocial behaviors in U.S. Latinx youth. There is also substantive support for positive links between ethnic identity (and bicultural identity) and U.S. Latinx youth prosocial behaviors (e.g., McGinley et al., 2020; Streit et al., 2018). This latter work suggests that U.S. Latinx youth who strongly adopt a strong ethnic or bicultural identity may be particularly motivated to engage in prosocial behaviors and that ethnic identity might encumber the traditional Latinx value of *bien educado*.

There is also evidence on the enhancing effects of ethnic socialization practices (see Carlo & Conejo, 2019). For example, Knight et al. (2016) demonstrated longitudinal empirical links between ethnic socialization practices and prosocial tendencies via youth endorsements of familism and ethnic identity. Similarly, in a recent study, Streit, Carlo, and Killoren (2020) showed evidence that parental ethnic socialization practices indirectly predicted prosocial behaviors through ethnic identity. These, and other studies (see Carlo & Conejo, 2019), support existing theories (Carlo & Conejo, 2019; Knight et al., 2015) of prosocial behaviors that assert the intervening role of ethnic identity in U.S. Latinx youth.

However, as ethnic/racial minorities in the U.S., there are also culture-related factors that enhance risk and can undermine U.S. Latinx youth adjustment (including prosocial behaviors). For example, discrimination experiences, including immigration status and language-based experiences, have been linked to lower levels of prosocial and altruistic behaviors (Brittian et al., 2013; Davis et al., 2016). In addition, acculturative stress, family conflict, and economic stress seem to have similar effects on U.S. Latinx youth prosocial behaviors (Davis, Carlo, Streit, & Crockett, 2018; McGinley et al., 2010; Streit, Carlo, Ispa, & Palermo, 2021).

2.4 Integration of traditional and cultural socialization theory and research

There is some research that has integrated traditional models of socialization and culture-specific socialization models in order to better understand parents' impact on prosocial development. For example, Streit and colleagues (2021) examined traditional parenting practices (i.e., acceptance and harsh parenting) and culture-group specific parenting practices (i.e., ethnic socialization) as predictors of multiple forms of prosocial behaviors via ethnic identity and familism values. In general, ethnic socialization practices predicted multiple forms of prosocial behaviors, including altruistic prosocial behaviors, via ethnic identity and familism (Streit, Carlo, Ispa, & Palermo, 2021; Streit, Carlo, Knight, White, & Maiya, 2021). In contrast, acceptance and harsh parenting directly predicted multiple forms of prosocial behaviors (Streit, Carlo, Ispa, & Palermo, 2021; Streit, Carlo, Knight, et al., 2021). This study illustrates the distinct predictive mechanisms of traditional parenting practices and culture-group specific parenting practices on U.S Latinx youth prosocial behaviors over time.

Taken together, socialization experiences, especially parenting and family influences, are intricately linked to prosocial development (see Carlo, 2014; Eisenberg et al., 2015) among both majority and minority groups. Therefore, such experiences are important to understand as they have the potential to combat the transmission of racism and prejudice across generations via the promotion of social justice orientations and prosocial actions towards diverse others.

3. Application of prosocial behaviors to address social injustice and inequities

3.1 A strengths-based approach to address social inequities and injustices

A recently posited model, *A U.S. Latinx Youth Model of Social Inequities*, highlighted prosocial behaviors as a key factor that might mitigate social inequities because of the impacts of prosocial behaviors on individuals and broader societal patterns (Davis et al., 2021). Overall, the evidence suggests that prosocial behaviors can be an avenue for promoting social well-being; intergroup harmony and cooperation among diverse populations and can reduce inequities by placing U.S. Latinx youth on a trajectory of social connection and integration. Prosocial behaviors can contribute to social integration, including deeper connections to community among youth as well as social engagement among U.S. Latinx youth (Frisco, Muller, & Dodson, 2004; Perez, Espinoza, Ramos, Coronado, & Cortes, 2010). Community connection, in turn, predicts indicators of social engagement and connection, including voting and connection to school (Frisco et al., 2004; Watson, Battistich, & Solomon, 1997). Moreover, promoting prosocial engagement can promote better health outcomes and reduces health inequities, including mental and physical health outcomes (Davis et al., 2016; Memmott-Elison, Holmgren, Padilla-Walker, & Hawkins, 2020; Schreier et al., 2013). In addition, researchers have demonstrated that prosocial behaviors can also buffer the negative effects of stress on physical health by elevating levels of neurotransmitters, such as oxytocin, that can promote physical health (Brown & Brown, 2015; Poulin & Holman, 2013).

While there is evidence that prosocial behaviors generally increase over the course of adolescents (Van der Graaff, Carlo, Crocetti, Koot, & Branje, 2018), there are also individual differences in prosocial behaviors that are important to consider. Understanding factors that foster prosocial behaviors is important, as there are a number of factors that can undermine these behaviors. For youth of color, there is evidence that discrimination and cultural stress experiences might inhibit altruistic behaviors (Brittian et al., 2013; Davis et al., 2016). Other forms of stress (e.g., economic stress and relational variables (e.g., family conflict, parental psychological control) have also been linked to lower levels of prosocial behaviors, including altruistic behaviors (Davis et al., 2018; Mouratidis, Sayil, Kumru, Selcuk, & Soenens, 2019). There are also sociocognitive and socioemotive variables that explain individual differences for diverse youth (e.g., moral reasoning, empathy; see Carlo, 2014; Taylor, O'Driscoll, Dautel, & McKeown, 2020) and these factors are particularly important to consider among majority youth as they help explain individual differences.

3.2 Predictors of prosociality between majority and minority groups

It is particularly important to promote prosocial behaviors toward social out-group members in order to promote group harmony and integration. The bulk of the research on prosocial behaviors has focused on types of helping, differentiated by helping in specific contexts and with differing motivations (see Carlo, 2014). However, helping behaviors also differ depending on target (Padilla-Walker & Christensen, 2011). It is critical to understand factors that promote prosocial behaviors toward cultural outgroup members in order to break down barriers and reduce "othering" processes that contribute to prejudice and discrimination. While evidence on the development of prosocial behaviors toward out-group members is scarce, there is some evidence that empathy and prosocial behaviors might differ depending on the characteristics of the target (Fabi & Leuthold, 2018; Yi, Todd, & Mekawi, 2020). Specifically, there is evidence that persons perceive light-skinned individuals as feeling more pain and also report more empathy toward light skinned individuals than dark-skinned individuals (Fabi & Leuthold, 2018), which suggests that racial bias is a driver of reduced empathic concern toward darker skinned individuals. Other work shows that color-blind racial attitudes were directly, negatively associated with intergroup empathy (Yi et al., 2020).

Discrimination experiences can also undermine prosocial behaviors, particularly altruistic prosocial behaviors (Brittian et al., 2013; Davis et al., 2016). U.S. Latinx adolescents who are experiencing discrimination may be drained of the cognitive and emotional resources that are necessary to engage in prosocial behaviors (see Batson & Powell, 2003; Lazarus & Folkman, 1984). These processes may lead to reduced levels of prosocial behaviors toward others, especially forms of helping primarily intended to benefit others with no expected benefit to the self (i.e., altruistic behaviors). There is evidence with Latinx youth, including recent immigrant youth, that discrimination experiences are negatively associated with altruistic prosocial behaviors (Brittian et al., 2013; Davis et al., 2016). Similarly, in a cross-sectional study of U.S. Mexican college students, McGinley et al. (2010) found that acculturative stress (which can include discrimination experiences and social stress associated with adapting to a new culture; Kulis, Marsiglia, & Nieri, 2009) was positively associated with multiple forms of prosocial behaviors but was negatively associated with altruistic prosocial behaviors. Discrimination and culture-related stressors, then, are a threat to undermine selflessly-motivated prosocial behaviors, which might be particularly important for majority-minority group members to cooperate and share resources more equitably.

While discrimination experiences might mitigate prosocial behaviors, sociocognitive and emotional characteristics can promote such behaviors. Specifically, prosocial moral reasoning and empathy are two characteristics that are salient predictors of multiple forms of prosocial behaviors (Fabes, Carlo, Kupanoff, & Laible, 1999). Empathic concern (i.e., feeling negative emotions consistent with the experience of another person; Hoffman, 2000) and prosocial moral reasoning may serve as indicators of moral internalization and intrinsic moral motivation and may be positively linked to multiple forms of prosocial behaviors. Perspective taking (i.e., cognitive component that reflects understanding the condition of another), on the other hand, is thought to be an important precursor to both empathic concern and prosocial moral reasoning (see Eisenberg & Fabes, 1990; Parker & Axtell, 2001). Research has consistently demonstrated links between prosocial moral reasoning and empathy, and multiple forms of prosocial behaviors (Carlo, Knight, et al., 2010; Carlo, Mestre, Samper, Tur, & Armenta, 2010; Eisenberg, 2002; see Carlo & de Guzman, 2009). Taken together, this work suggests avenues for fostering greater prosociality and altruistic tendencies, which can promote positive intergroup relationships.

3.3 Implications for interventions aimed at addressing social injustices

Prosocial behaviors include micro-level actions (helping specific persons) and also broader, community-focused actions (e.g., civic engagement). Civic engagement includes behaviors such as volunteering, voting, community activism, and involvement in groups that work to better the community (Keeter, Zukin, Andolina, & Jenkins, 2002). Understanding factors that promote prosocial behaviors aimed at broader social change are important because such actions can result in structural and systemic changes necessary to modify the legacy of bias and prejudice towards minorities. However, such change is founded on agency in persons who are willing to act prosocially towards out-group members. Majority group members must commit themselves to engage in prosocial actions towards minority group members to reduce social injustice. At the same time, however, minority group members must be willing to reciprocally engage in prosocial actions with majority group members to reduce social marginalization and isolation. Thus, the need for interventions that foster prosociality among both majority and minority group members is necessary for successful intergroup harmony and integration for minoritized youth, including U.S. Latinx youth.

Although there is great need for further developmental research on ingroup/outgroup prosocial behaviors, there is promising evidence for intervention approaches that could address social inequities. For example, the work on sociocognitive and socioemotive mechanisms (e.g., prosocial moral reasoning, sympathy) suggests intervention points to develop programs that increase prosocial tendencies, including altruistic prosocial behaviors. Other scholarly work on parenting shows promise for parenting education programs that could encourage the use of authoritative and inductive parenting practices, prosocial parenting practices (especially the use of social rewards), and proactive parenting practices to foster youth prosocial behaviors. Researchers have also identified factors that can mitigate prosocial behaviors including stress (e.g., economic stress), family conflict, deviant peer affiliation, and harsh parenting practices. Among U.S. Latinx youth, programs that strengthen and support ethnic socialization practices, cultural values (e.g., familism), and ethnic identity could facilitate prosocial actions, which could result in greater social integration and cooperation. In contrast, efforts that teach youth to manage or avoid exposure to discrimination and cultural stress could also enhance prosocial tendencies.

The challenges of addressing the long-standing and enduring social injustices and inequities in our systems are great. The research on prosocial development provides promise in identifying foundational mechanisms of positive change at the individual level that can translate into positive change at the interpersonal, group, and societal levels. Creating reciprocated prosocial actions between majority and minority group members is a daunting task. Successful intervention likely requires prevention and intervention efforts that can disrupt the pattern of selfishly motivated and harmful attitudes, and actions towards outgroup members that can be acquired early in life. Although interventions are needed to redress antisocial attitudes and behaviors, intervention programs in early childhood will be of utmost importance to foster early prosocial tendencies towards diverse others and to mitigate antisocial development. These comprehensive efforts that are designed to promote prosocial behaviors towards all are much needed to help redress the historical systemic biases and inequities that plague our societies.

References

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Batson, C. D. (1998). Altruism and prosocial behavior. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.) (Vols. 1 and 2). *The handbook of social psychology* (pp. 282–316). New York, NY: McGraw-Hill.

Batson, C. D., & Powell, A. A. (2003). Altruism and prosocial behavior. In *Handbook of psychology*. Hoboken, NJ: John Wiley and Sons. https://doi.org/10.1002/0471264385.wei0519.

Berlan, M., & Harwood, M. (2018). *National snapshot: Poverty among women and families, 2018*. The National Women's Law Center. https://nwlc.org/wp-content/uploads/2018/09/National-Snapshot.pdf.

Berry, J. W. (2017). Theories and models of acculturation. In S. J. Schwartz, & J. Unger (Eds.), *The Oxford handbook of acculturation and health*. Oxford University Press.

Bridges, M., Cohen, S. R., McGuire, L. W., Yamada, H., Fuller, B., Mireles, L., et al. (2012). Bien educado: Measuring the social behaviors of Mexican American children. *Early Childhood Research Quarterly*, *27*(3), 555–567. https://doi.org/10.1016/j.ecresq.2012.01.005.

Brittian, A. S., O'Donnell, M., Knight, G. P., Carlo, G., Umaña-Taylor, A. J., & Roosa, M. W. (2013). Associations between adolescents' perceived discrimination and prosocial tendencies: The mediating role of Mexican American values. *Journal of Youth and Adolescence*, *42*(3), 328–341. https://doi.org/10.1007/s10964-012-9856-6.

Brown, S. L., & Brown, R. M. (2015). Connecting prosocial behavior to improved physical health: Contributions from the neurobiology of parenting. *Neuroscience & Biobehavioral Reviews*, *55*, 1–17. https://doi.org/10.1016/j.neubiorev.2015.04.004.

Carlo, G. (2014). The development and correlates of prosocial moral behaviors. In M. Killen, & J. G. Smetana (Eds.), *Handbook of moral development* (2nd ed.). New York, NY: Psychology Press.

Carlo, G., & Conejo, L. D. (2019). Traditional and culture-specific parenting of prosociality in U.S. Latino/as. In D. Laible, L. Padilla-Walker, & G. Carlo (Eds.), *Oxford handbook of parenting and moral development*. New York: Oxford University Press.

Carlo, G., & de Guzman, M. R. T. (2009). Theories and research on prosocial competencies among US Latinos/as. In F. Villaruel, G. Carlo, M. Azmitia, J. Grau, N. Cabrera, & J. Chahin (Eds.), *Handbook of U.S. Latino Psychology* (pp. 191–211). Sage Publications.

Carlo, G., Knight, G. P., McGinley, M., & Hayes, R. (2011). The roles of parental inductions, moral emotions, and moral cognitions in prosocial tendencies among Mexican American and European American early adolescents. *Journal of Early Adolescence*, *31*, 757–781. https://doi.org/10.1177/0272431610373100.

Carlo, G., Knight, G. P., McGinley, M., Zamboanga, B. L., & Jarvis, L. H. (2010). The multidimensionality of prosocial behaviors and evidence of measurement equivalence in Mexican American and European American early adolescents. *Journal of Research on Adolescence*, *20*, 334–358. https://doi.org/10.1111/j.1532-7795.2010.00637.x.

Carlo, G., McGinley, M., Hayes, R. C., & Martinez, M. M. (2012). Empathy as a mediator of the relations between parent and peer attachment and prosocial and physically aggressive behaviors in Mexican American college students. *Journal of Social and Personal Relationships*, *29*(3), 337–357. https://doi.org/10.1177/0265407511431181.

Carlo, G., Mestre, M. V., McGinley, M. M., Tur-Porcar, A., Samper, P., & Opal, D. (2014). The protective role of prosocial behaviors on antisocial behaviors: The mediating effects of deviant peer affiliation. *Journal of Adolescence*, *37*(4), 359–366. https://doi.org/10.1016/j.adolescence.2014.02.009.

Carlo, G., Mestre, M. V., Samper, P., Tur, A., & Armenta, B. E. (2010). Feelings or cognitions? Moral cognitions and emotions as longitudinal predictors of prosocial and aggressive behaviors. *Personality and Individual Differences*, *48*(8), 872–877. https://doi.org/10.1016/j.paid.2010.02.010.

Carlo, G., & Randall, B. A. (2002). The development of a measure of prosocial behaviors for late adolescents. *Journal of Youth and Adolescence*, *31*(1), 31–44. https://doi.org/10.1023/A:1014033032440.

Carlo, G., White, R. M., Streit, C., Knight, G. P., & Zeiders, K. H. (2018). Longitudinal relations among parenting styles, prosocial behaviors, and academic outcomes in US Mexican adolescents. *Child Development*, *89*(2), 577–592. https://doi.org/10.1111/cdev.12761.

Davis, A. N., & Carlo, G. (2018). The roles of parenting practices, sociocognitive/emotive traits, and prosocial behaviors in low-income adolescents. *Journal of Adolescence*, *62*, 140–150. https://doi.org/10.1016/j.adolescence.2017.11.011.

Davis, A. N., Carlo, G., Hardy, S. A., Olthuis, J. V., & Zamboanga, B. L. (2017). Bidirectional relations between different forms of prosocial behaviors and substance use among female college student athletes. *The Journal of Social Psychology*, *6*(157), 645–657. https://doi.org/10.1080/00224545.2016.1263596.

Davis, A. N., Carlo, G., & Maiya, S. (2021). Towards a multisystems, strengths-based model of social inequities in U.S. latinx youth. *Human Development*, *65*, 204–216. https://doi.10.1159/000517920.

Davis, A. N., Carlo, G., Schwartz, S. J., Unger, J. B., Zamboanga, B. L., Lorenzo-Blanco, E. I., et al. (2016). The longitudinal associations between discrimination, depressive symptoms, and prosocial behaviors in US Latino/a recent immigrant adolescents. *Journal of Youth and Adolescence*, *45*, 457–470. https://doi.org/10.1007/s10964-015-0394-x.

Davis, A. N., Carlo, G., Streit, C., & Crockett, L. J. (2018). Considering economic stress and empathic traits in predicting prosocial behaviors among U.S. Latino adolescents. *Social Development*, *27*(1), 58–72. https://doi.org/10.1111/sode.12249.

de Guzman, M. R. T., Do, K. A., & Kok, C. M. (2014). The cultural contexts of children's prosocial behaviors. In L. M. Padilla-Walker, & G. Carlo (Eds.), *Prosocial development: A multidimensional approach*. Oxford University Press.

Doll, J. J., Eslami, Z., & Walters, L. (2013). Understanding why students drop out of high school, according to their own reports: Are they pushed or pulled, or do they fall out? A comparative analysis of seven nationally representative studies. *Sage Open*, *3*(4), 1–15. https://doi.org/10.1177/2158244013503834.

Eisenberg, N. (1986). *Altruistic cognition, emotion, and behavior*. Hillsdale, NJ: Erlbaum.

Eisenberg, N. (2002). Empathy-related emotional responses, altruism, and their socialization (2002). In R. J. Davidson, & A. Harrington (Eds.), *Visions of compassion: Western scientists and Tibetan Buddhists examine human nature* (pp. 131–164). New York, NY: Oxford.

Eisenberg, N., Eggum-Wilkens, N. D., & Spinrad, T. L. (2015). The development of prosocial behavior. In D. A. Schroeder, & W. G. Graziano (Eds.), *The Oxford handbook of prosocial behavior* (pp. 114–136). Oxford, UK: Oxford University Press.

Eisenberg, N., & Fabes, R. A. (1990). Empathy: Conceptualization, measurement, and relation to prosocial behavior. *Motivation and Emotion*, *14*(2), 131–149. https://doi.org/10.1007/BF00991640.

Eisenberg, N., & Fabes, R. A. (1991). Prosocial behavior and empathy: A multimethod developmental perspective. In M. S. Clark (Ed.), *Prosocial behavior* (pp. 34–61). Newbury Park, CA: Sage.

Eisenberg, N., Fabes, R. A., & Spinard, T. L. (2006). Prosocial development. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Social, emotional and personality development: Vol. 3. Handbook of child psychology* (6th ed., pp. 646–718). New York: John Wiley.

Espinola, M., Zhen-Duan, J., Suarez-Cano, G., Mowry-Mora, I., & Schultz, J. M. (2019). The impact of US sociopolitical issues on the prejudicial treatment of Latino children and youth. In H. Fitzgerald, D. Johnson, D. Qin, F. Villarruel, & J. Norder (Eds.), *Handbook of children and prejudice: Integrating research, practice, and policy*. Cham, Switzerland: Springer Press.

Fabes, R. A., Carlo, G., Kupanoff, K., & Laible, D. (1999). Early adolescence and prosocial/moral behavior I: The role of individual processes. *The Journal of Early Adolescence*, *19*(1), 5–16. https://doi.org/10.1177/0272431699019001001.

Fabi, S., & Leuthold, H. (2018). Racial bias in empathy: Do we process dark-and fair-colored hands in pain differently? An EEG study. *Neuropsychologia*, *114*, 143–157. https://doi.org/10.1016/j.neuropsychologia.2018.04.024.

Federal Bureau of Investigation. (2020). *FBI urges vigilance during COVID-19 pandemic*. Retrieved from https://www.fbi.gov/coronavirus.

Flores, E., Tschann, J. M., Dimas, J. M., Bachen, E. A., Pasch, L. A., & de Groat, C. L. (2008). Perceived discrimination, perceived stress, and mental and physical health among Mexican-origin adults. *Hispanic Journal of Behavioral Sciences*, *30*, 401–424. https://doi.org/10.1177/0739986308323056.

Frisco, M., Muller, C., & Dodson, K. (2004). Participation in voluntary youth-serving associations and early adult voting behavior. *Social Science Quarterly*, *85*, 660–676. https://doi.org/10.1111/j.0038-4941.2004.00238.x.

García-Coll, C., Lamberty, G., Jenkins, R., Pipes McAdoo, H. P., Crnic, K., Wasik, B. H., et al. (1996). An integrative model for the study of developmental competencies in minority children. *Child Development*, *67*, 1891–1914. https://doi.org/10.1111/j.1467-8624.1996.tb01834.x.

Gülseven, Z., & Carlo, G. (2021). Parenting and prosocial behaviors in Nicaraguan adolescents: The roles of prosocial moral reasoning and familism. *Journal of Social and Personal Relationships*, *02654075211009298*.

Hoffman, M. L. (2000). *Empathy and moral development: Implications for caring and justice*. Cambridge, UK: Cambridge University Press.

Keeter, S., Zukin, C., Andolina, M., & Jenkins, K. (2002). *The civic and political health of the nation: A generational portrait*. Center for Information and Research on Civic Learning and Engagement (CIRCLE).

Knight, G. P., Bernal, M. E., Garza, C. A., & Cota, M. K. (1993). A social cognitive model of ethnic identity and ethnically-based behaviors. In M. E. Bernal, & G. P. Knight (Eds.), *Ethnic identity: Formation and transmission among Hispanics and other minorities* (pp. 213–234). Albany, NY: State University of New York Press.

Knight, G. P., & Carlo, G. (2012). Prosocial development among Mexican American youth. *Child Development Perspectives*, *6*(3), 258–263. https://doi.org/10.1111/j.1750-8606.2012.00233.x.

Knight, G. P., Carlo, G., Basilio, C. D., & Jacobson, R. P. (2015). Familism values, perspective taking, and prosocial moral reasoning: Predicting prosocial tendencies among Mexican American adolescents. *Journal of Research on Adolescence*, *25*(4), 717–727. https://doi.org/10.1111/jora.12164.

Knight, G. P., Carlo, G., Mahrer, N. E., & Davis, A. N. (2016). The socialization of culturally related values and prosocial tendencies among Mexican American adolescents. *Child Development*, *87*(6), 1758–1771. https://doi.org/10.1111/cdev.12634.

Knight, G. P., Gonzales, N. A., Saenz, D. S., Bonds, D. D., Germán, M., Deardorff, J., et al. (2010). The Mexican American cultural values scale for adolescents and adults. *The Journal of Early Adolescence*, *30*(3), 444–481. https://doi.org/10.1177/0272431609338178.

Knight, G. P., Mazza, G. L., & Carlo, G. (2018). Trajectories of familism values and the prosocial tendencies of Mexican American adolescents. *Developmental Psychology*, *54*, 378–384. https://doi.org/10.1037/dev0000436.

Krevans, J., & Gibbs, J. C. (1996). Parents' use of inductive discipline: Relations to children's empathy and prosocial behavior. *Child Development*, *67*(6), 3263–3277. https://doi.org/10.1111/j.1467-8624.1996.tb01913.x.

Kulis, S., Marsiglia, F. F., & Nieri, T. (2009). Perceived ethnic discrimination versus acculturation stress: Influences on substance use among Latino youth in the Southwest. *Journal of Health and Social Behavior*, *50*, 443–459. https://doi.org/10.1177/002214650905000405.

Kumru, A., Carlo, G., Mestre, M. V., & Samper, P. (2012). Prosocial moral reasoning and prosocial behavior among Turkish and Spanish adolescents. *Social Behavior and Personality: An International Journal*, *40*(2), 205–214. https://doi.org/10.2224/sbp.2012.40.2.205.

Laosa, L. M., & Henderson, R. W. (1991). Cognitive socialization and competence: The academic development of Chicanos. *ETS Research Report Series*, *1991*(2). i-77 https://doi.org/10.1002/j.2333-8504.1991.tb01406.x.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer.

McGinley, M. (2018). Can hovering hinder helping? Examining the joint effects of helicopter parenting and attachment on prosocial behaviors and empathy in emerging adults. *The Journal of Genetic Psychology*, *179*(2), 102–115.

McGinley, M., Carlo, G., Raffaelli, M., Crockett, L., Torres-Stone, R., & Itubirde, M. (2010). Stressed and helping: The relations among acculturative stress, gender, and prosocial behaviors in Mexican Americans. *Journal of Social Psychology*, *150*, 34–56. https://doi.org/10.1080/00221325.2018.1438985.

McGinley, M., Davis, A. N., Carlo, G., Schwartz, S. J., Lorenzo-Blanco, E. I., Unger, J. B., et al. (2020). A parallel process model of integration and multidimensional prosocial behaviors in recent immigrant U.S. Latinx adolescents. *Psychological Reports*. https://doi.org/10.1177/0033294120928268.

Memmott-Elison, M. K., Holmgren, H. G., Padilla-Walker, L. M., & Hawkins, A. J. (2020). Associations between prosocial behavior, externalizing behaviors, and internalizing symptoms during adolescence: A meta-analysis. *Journal of Adolescence*, *80*, 98–114. https://doi.org/10.1016/j.adolescence.2020.01.012.

Mouratidis, A. A., Sayil, M., Kumru, A., Selcuk, B., & Soenens, B. (2019). Maternal knowledge as a mediator of the relation between maternal psychological control and altruistic prosocial, instrumental prosocial, and antisocial behavior. *Merrill-Palmer Quarterly*, *65*, 207–231. https://doi.org/10.13110/merrpalmquar1982.65.2.0207.

Padilla-Walker, L. M., & Christensen, K. J. (2011). Empathy and self-regulation as mediators between parenting and adolescents' prosocial behavior toward strangers, friends, and family. *Journal of Research on Adolescence*, *21*(3), 545–551. https://doi.org/10.1111/j.1532-7795.2010.00695.x.

Padilla-Walker, L. M., Fraser, A. M., & Harper, J. M. (2012). Walking the walk: The moderating role of proactive parenting on adolescents' value-congruent behaviors. *Journal of Adolescence*, *35*(5), 1141–1152. https://doi.org/10.1016/j.adolescence.2012.03.003.

Parker, S. K., & Axtell, C. M. (2001). Seeing another viewpoint: Antecedents and outcomes of employee perspective taking. *Academy of Management Journal*, *44*(6), 1085–1100. https://doi.org/10.5465/3069390.

Perez, W., Espinoza, R., Ramos, K., Coronado, H., & Cortes, R. (2010). Civic engagement patterns of undocumented Mexican students. *Journal of Hispanic Higher Education*, *9*, 245–265. https://doi.org/10.1177/1538192710371007.

Poulin, M. J., & Holman, E. A. (2013). Helping hands, healthy body? Oxytocin receptor gene and prosocial behavior interact to buffer the association between stress and physical health. *Hormones and Behavior*, *63*(3), 510–517. https://doi.org/10.1016/j.yhbeh.2013.01.004.

Raffaelli, M., Carlo, G., Carranza, M. A., & Gonzales-Kruger, G. E. (2005). Understanding Latino children and adolescents in the mainstream: Placing culture at the center of developmental model. In R. Larson, & Jensen (Eds.), *New horizons in developmental research: New directions for child and adolescent development* (pp. 23–32). San Francisco: Jossey-Bass.

Schreier, H. M., Schonert-Reichl, K. A., & Chen, E. (2013). Effect of volunteering on risk factors for cardiovascular disease in adolescents: A randomized controlled trial. *JAMA Pediatrics*, *167*(4), 327–332. https://doi.org/10.1001/jamapediatrics.2013.1100.

Spitzer, U. S., & Hollmann, W. (2013). Experimental observations of the effects of physical exercise on attention, academic and prosocial performance in school settings. *Trends in Neuroscience & Education*, *2*(1), 1–6. https://doi.org/10.1016/j.tine.2013.03.002.

Staub, E. (1978). Positive social behavior and morality. In *Vol. 1. Social and personal influences*. New York: Academic Press.

Staub, E. (2005). The roots of goodness: The fulfillment of basic human needs and the development of caring. In G. Carlo, & C. P. Edwards (Eds.), *Vol. 51. Nebraska Symposium on Motivation* (pp. 33–72). Lincoln, NE: University of Nebraska.

Streit, C., Carlo, G., Ispa, J., & Palermo, F. (2021). Direct and indirect relations between family conflict and youth's later behavioral outcomes. *Social Development*, *30*, 482–495. https://doi.org/10.1111/sode.12483.

Streit, C., Carlo, G., & Killoren, S. E. (2020). Ethnic socialization, identity, and values associated with US Latino/a young adults' prosocial behaviors. *Cultural Diversity and Ethnic Minority Psychology*, *26*(1), 102–111. https://doi.org/10.1037/cdp0000280.

Streit, C., Carlo, G., Killoren, S. E., & Alfaro, E. C. (2018). Family members' relationship qualities and prosocial behaviors in US Mexican young adults: The roles of familism and ethnic identity resolution. *Journal of Family Issues*, *39*(4), 1056–1084. https://doi.org/10.1177/0192513X16686134.

Streit, C., Carlo, G., Knight, G. P., White, R. M. B., & Maiya, S. (2021). Relations among parenting, culture, and prosocial behaviors in U.S. Mexican youth: An integrative socialization approach. *Child Development*. https://doi.org/10.1111/cdev.13550.

Super, C., & Harkness, S. (1997). The cultural structuring of child development. In J. Berry, P. Dasen, & T. Saraswathi (Eds.), *Basic process and human development: Vol. 2. Handbook of cross-cultural psychology* (2nd ed., pp. 1–39). Boston: Allyn and Bacon.

Taylor, L. K. (2020). The developmental peacebuilding model (DPM) of children's prosocial behaviors in settings of intergroup conflict. *Child Development Perspectives*, *14*(3), 127–134. https://doi.org/10.1111/cdep.12377.

Taylor, L. K., O'Driscoll, D., Dautel, J. B., & McKeown, S. (2020). Empathy to action: Child and adolescent out-group attitudes and prosocial behaviors in a setting of intergroup conflict. *Social Development*, *29*(2), 461–477. https://doi.org/10.1111/sode.12421.

Van der Graaff, J., Carlo, G., Crocetti, E., Koot, H. M., & Branje, S. (2018). Prosocial behavior in adolescence: Gender differences in development and links with empathy. *Journal of Youth and Adolescence*, *47*(5), 1086–1099. https://doi.org/10.1007/s10964-017-0786-1.

Walters, G. D. (2020). Prosocial peers as risk, protective, and promotive factors for the prevention of delinquency and drug use. *Journal of Youth and Adolescence*, *49*(3), 618–630. https://doi.org/10.1007/s10964-019-01058-3.

Watson, M., Battistich, V., & Solomon, D. (1997). Enhancing students' social and ethnical development in school: An intervention program and its effects. *International Journal of Educational Research*, *27*, 571–586. https://doi.org/10.1016/S0883-0355(97)00055-4.

Whiting, B. B., & Whiting, J. W. M. (1975). *Children of six cultures: A psycho-cultural analysis*. Cambridge, MA: Harvard University Press.

Yi, J., Todd, N. R., & Mekawi, Y. (2020). Racial colorblindness and confidence in and likelihood of action to address prejudice. *American Journal of Community Psychology*, *65*(3-4), 407–422. https://doi.org/10.1002/ajcp.12409.

CHAPTER SIX

Pathways for engaging in prosocial behavior in adolescence

Eveline A. Crone[a,b,*], Sophie W. Sweijen[a], Lysanne W. te Brinke[a], and Suzanne van de Groep[a]

[a]Erasmus School of Social and Behavioral Sciences, Erasmus University Rotterdam, Rotterdam, The Netherlands
[b]Department of Developmental and Educational Psychology, Leiden University, Leiden, The Netherlands
*Corresponding author: e-mail address: crone@essb.eur.nl

Contents

Abstract

Adolescent development is often regarded as a period of social sensitivities, given that brain development continues into the early 20s in interplay with social experiences. In this review, we present adolescence as a unique window for prosocial development; that is, behavior that benefits others. We present evidence for multiple pathways of neural sensitivity that contribute to key developmental processes related to prosocial behaviors, including valuing, perspective taking, and goal-flexibility. Yet, these processes are dependent on several contextual factors including recipients, audience effects, and strategic motivations. Next, we present intervention findings suggesting that prosocial

Advances in Child Development and Behavior, Volume 63
ISSN 0065-2407
https://doi.org/10.1016/bs.acdb.2022.03.003

experiences within these various contexts are crucial for adolescents developing into engaged and contributing members of society. These findings suggest a new interpretation of the elevated socio-affective sensitivity and emerging socio-cognitive development in adolescence, focusing on opportunities rather than risks.

1. Introduction

One of the hallmarks of adolescence is developing meaningful relationships outside the family context to eventually become an engaged and contributing member of society (Fuligni, 2019). As such, adolescence is an important transition period between the dependency on parents and other caregivers of childhood, and the mature social goals and independence of adulthood. The development of prosocial behavior, defined as behavior intended to benefit others, is of crucial importance for taking social responsibilities and developing mature social relationships (Carlo & Padilla-Walker, 2020). These behaviors may lead to extension of relationships with family members and friends to contributing to needs of more distant others (e.g., helping unknown others) and to society (e.g., engaging in community service).

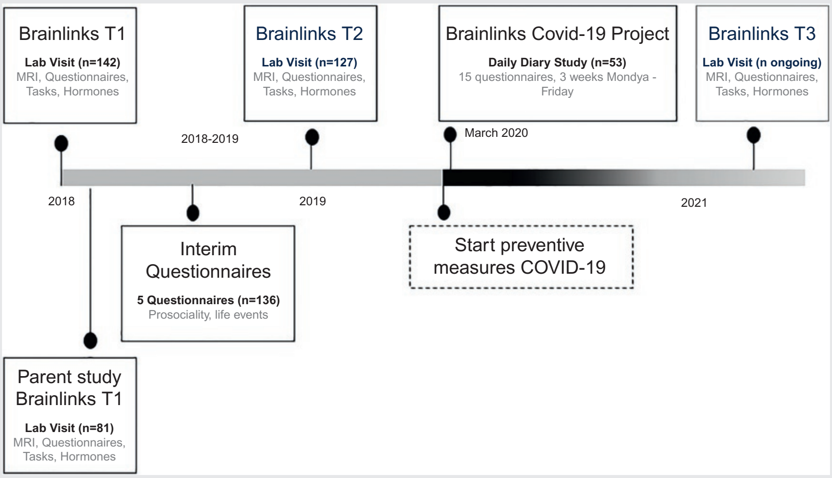
Adolescence represents a developmental time window typified by strong needs for exploration, forming new relationships, increasing intimacy, and rapid adjustment to changing social contexts (Blakemore & Mills, 2014; Steinberg, 2008). Adolescence starts with the onset of puberty, approximately at the age of 9–10-years in girls and 10–11-years in boys, although differences are observed between countries and cultures (Crone & Dahl, 2012). The onset of puberty is characterized by a rise in gonadal hormones, which are released through the hypothalamus-pituitary-gonadal axis, and have large influence on bodily characteristics and brain development (Goddings et al., 2014; Shirtcliff, Dahl, & Pollak, 2009). Pubertal development (also referred to as early to mid-adolescence) lasts until approximately age 15–16-years (with differences between cultures). The prolonged period of mid- to late adolescence continues until individuals have achieved mature social and personal responsibilities and is culturally dependent (Steinberg, 2008). The period of adolescence has extended considerably in the last century, as individuals rely on parents longer and have more possibilities for personal development and identity formation (Arnett, Zukauskiene, & Sugimura, 2014).

Prosocial behavior refers to behavior that benefits others, which can be non-costly (e.g., helping) or costly (e.g., sharing) (Carlo & Padilla-Walker, 2020). It is well established that prosocial behavior changes during adolescence and young adulthood. However, the exact developmental patterns are still debated. Some studies show increases in prosocial behavior during adolescence (Fu, Padilla-Walker, & Brown, 2017; Padilla-Walker, Carlo, & Memmott-Elison, 2018), whereas others find decreases or stabilization (Malti et al., 2015). This has recently been interpreted as evidence that prosocial behavior should be regarded as a multi-dimensional construct, comprising many behaviors such as cooperating, helping, and giving. It depends on socio-affective and socio-cognitive developmental processes, (i.e., valuing, perspective-taking, and goal flexibility) and on multiple contextual processes (i.e., the target or recipient of prosocial behavior and the situational context), such as whether prosocial behavior is observed by others (i.e., audience effects; Carlo & Padilla-Walker, 2020).

In this review article, we will provide a comprehensive perspective on prosocial development suggesting that the development of prosocial behavior depends on mutual socio-affective and socio-cognitive maturation, as is evident from behavioral and neural pathways. The review is organized around the common themes of the Brainlinks study, an experimental accelerated longitudinal cohort and intervention study on prosocial development that includes adolescents between ages 9–22-years (see Box 1 for a description and Appendix 1 for the meta-data). First, we will summarize evidence for the structural development of the human brain during adolescence, suggesting that this may be a time during which the developing individual is particularly sensitive to environmental influences. Second, we will present a possible model to describe the pathways of multi-dimensionality of prosocial behavior, illustrated with examples of recent empirical developmental behavioral and neuroimaging studies (including, but not limited to results from the Brainlinks study). Third, we will show behavioral evidence for malleability of prosocial behavior in intervention designs according to the presented model. Finally, we will show that a comprehensive understanding of pathways of prosocial behavior, including sensitivity to environmental influences, will be of importance for valuing the contribution of young people to benefit self and other, as well as to adaptation and resilience of the society at large.

BOX 1 The Brainlinks study

The Brainlinks study located in the Netherlands aims to examine the multidimensionality of prosocial development in adolescence. The project includes (a) an accelerated longitudinal design with neuroimaging, behavioral, and questionnaire measures acquired in 3 waves across a time window of 5 years (see figure); and (b) a behavioral intervention study. We present the full project's meta-data in the Appendix 1.



The goal of the longitudinal study is to examine the processes described in this review (vicarious gains/cooperation, giving/sharing, trust/reciprocity), which reflect prosocial processes of increasing complexity. Contextual manipulations involve the target of prosocial behavior, audience effects on giving and strategic giving manipulations. The key aims of the project are:

(i) Combining neural development with behavioral development
(ii) A longitudinal design allowing for testing within-person change
(iii) A multidimensional approach of prosocial behavior allowing for the examination of state and trait dimensions, and possible underlying latent variables, as well as associations within individuals over time
(iv) A detailed assessment of environmental factors that may shape prosocial development

Prior studies have pointed to an important role of the family in shaping prosocial behaviors in adolescence. The study therefore includes an enrichment wave (following wave 1) in which a selection of mothers and fathers performed a selection of the same measures as the adolescents, including a neuroimaging vicarious reward task for children and a self-concept task for parenting.

BOX 1 The Brainlinks study—cont'd

An additional aim of the study includes a separate naturalistic and micro-trial intervention study for enrichment of prosocial experiences, with partly overlapping measures as the longitudinal study. These study dimensions will reveal the effects of environmental factors hindering or fostering prosocial development. An unexpected event was the start of the pandemic between wave 2 and wave 3 of the longitudinal study. For this reason, the study has an enrichment of three-weeks daily diary measures collected in between these waves.

2. A cognitive neuroscience perspective on adolescent development

By the time children enter puberty, the human brain has already gone through massive developments (Gilmore, Knickmeyer, & Gao, 2018; Lee et al., 2019). Yet, numerous longitudinal structural neuroimaging studies have revealed that adolescent development involves additional widespread changes in the structure of the brain (Mills et al., 2016; Tamnes et al., 2017). Longitudinal research examining changes in brain structure over time within individuals has shown that cortical white matter increases approximately linearly with age throughout childhood and adolescence until the early twenties (Paus, 2010). In addition, cortical gray matter, which reflects neuronal density and the number of connections between neurons, follows an inverted-U shape over development, peaking at different ages depending on the region (Tamnes et al., 2017). Therefore, gray matter loss is often considered an index of the time-course of maturation of a brain region (Lee et al., 2014).

Within the cortex, gray matter reduction is most protracted for medial and lateral prefrontal cortex (PFC), and the junction between temporal cortex and parietal cortex (temporal-parietal junction: TPJ). Here, cortical gray matter loss continues until the early 20s (Mills, Goddings, Clasen, Giedd, & Blakemore, 2014). The development of subcortical brain regions, which are evolutionarily older parts of the brain, is also subject to both linear and non-linear changes, such that some subcortical regions (such as the caudate and the putamen) linearly decrease in size, whereas other subcortical regions (such as the amygdala and the hippocampus) show an increase in size at the onset of puberty, which stabilizes in adolescence and adulthood (Herting et al., 2018; Wierenga et al., 2018). Both cortical and subcortical brain development are driven by both age- and puberty-specific changes

(Goddings et al., 2014). The relation between structural brain volume changes and changes in behavior, however, is currently not yet well understood. In addition, very little is known about how experience-dependent changes influence or shape brain development in adolescence. Recent studies show initial evidence for an important contribution of social experiences on brain development, by showing a longitudinal relationship between gray matter thickness of the medial PFC and TPJ, and friendship quality (Becht et al., 2020). More direct evidence for the hypothesis that brain development is sensitive to environmental influences comes from genetic twin modeling. In a recent study in 7–8-year-old monozygotic and dizygotic twins, it was found that gray matter of all social brain regions is heritable, but that especially the gray matter thickness of TPJ was sensitive to shared environmental influences (Van der Meulen et al., 2020). Taken together, structural brain imaging findings illustrate a formative change in gray matter thickness and surface area during childhood and the teenage years, with initial evidence that some of the regions that show the most protracted development are more sensitive to environmental influences.

One way to further understand the relation between brain development and behavior is by using event-related functional magnetic resonance imaging (fMRI) studies. fMRI gives insight into neural regions that are involved in processing events or decision-making by detecting changes in blood oxygenation and flow that occur in response to neural activity (Logothetis, 2008). Known for its excellent spatial resolution, fMRI enables the examination of both cortical and subcortical brain regions, which are both assumed to play an important role in social behaviors (Blakemore & Mills, 2014; Crone & Dahl, 2012; Logothetis, 2008). An additional advantage of fMRI is that it allows for the measurement of processes that may be hard to capture on a behavioral level, such as initial tendencies, feelings, and other processes that are not necessarily expressed verbally or behaviorally (Lieberman, Straccia, Meyer, Du, & Tan, 2019). To date, studies have reported separate developmental pathways for socio-affective processes, with a focus on the subcortical ventral striatum (VS), and socio-cognitive developmental processes, with a focus on cortical brain regions including the medial PFC, the TPJ, and the superior temporal sulcus (STS), and the lateral prefrontal cortex (DLPFC) (Fig. 1). We will summarize these processes below in a pathway model, followed by empirical examples.

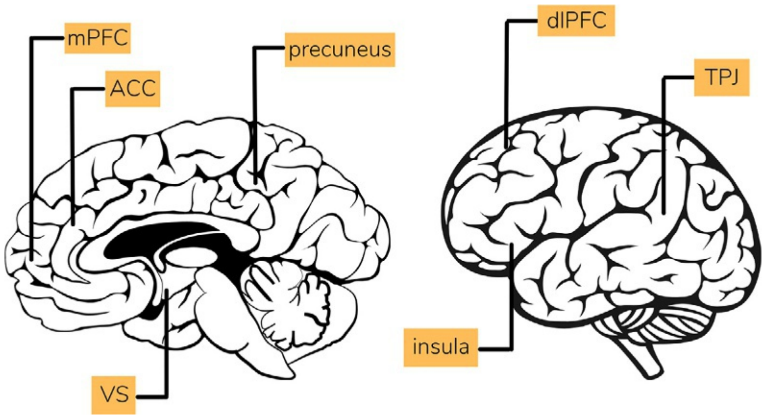


Fig. 1 Brain regions involved in various aspects of prosocial behavior, displayed are the ventral striatum (VS), insula, anterior cingulate cortex (ACC), medial prefrontal cortex (mPFC), precuneus, temporal parietal junction (TPJ), the superior temporal sulcus (STS) and the dorsolateral prefrontal cortex (dlPFC).

3. Capturing the complexity of prosocial development in a multiple-pathway model

The complexity of prosocial development can best be understood by decomposition of the various processes involved in behaviors that benefit self and others (Tamir & Hughes, 2018). Here, we differentiate between socio-affective and socio-cognitive processes which are thought to follow separable developmental time courses. Adolescence is one of the periods in life well known for its rise in emotional reactivity, both in terms of frequency and intensity (Dahl, Allen, Wilbrecht, & Suleiman, 2018). The dynamic characteristic of emotional reactivity is thought to peak in mid-adolescence, suggesting that adolescents can experience emotions more strongly than children and adults (Ernst, 2014; Larson, Moneta, Richards, & Wilson, 2002).

These changes in emotional reactivity and reward processing co-occur with a protracted development of socio-cognitive perspective taking. Whereas it has been well conceptualized that the basic socio-cognitive building blocks for prosocial behavior, such as theory of mind, develop in early childhood, recent studies have supported the notion that more complex social-cognitive behaviors, such as perspective taking and goal-flexibility, emerge in adolescence (Dumontheil, Apperly, & Blakemore, 2010; Fett et al., 2014; van den Bos, Westenberg, Van Dijk, & Crone, 2010).

Traditionally, the heightened emotional reactivity and protracted development of socio-cognitive functions have been linked to maladaptive adolescent behaviors such as alcohol and substance abuse, anxiety, and depression (Paus, Keshavan, & Giedd, 2008). These behaviors are known to increase considerably in adolescence with peaks in risk-taking and social anxiety around age 16–17-years (Blote, Kint, Miers, & Westenberg, 2009). This is also the time when most affect-driven psychiatric disorders manifest themselves for the first time, such as anxiety, depression, substance abuse, and schizophrenia (Lee et al., 2014; Paus et al., 2008).

Existing models, however, have often ignored how this normative development of emotional reactivity and protracted development of socio-cognitive functions have adaptive functions, creating opportunities to understand people's views and motivations (i.e., perspective taking) and aiding rapid adaptation to different contexts (i.e., goal flexibility; (Crone & Dahl, 2012)). Recent evidence suggests that the very same emotional reactivity that creates sensitivities for potential negative developmental trajectories (including risk for substance abuse, delinquency, social anxiety, or depression) may under other circumstances create opportunities for positive developmental trajectories—such as by fostering social sensitivity, cooperation, sharing, and helping (Telzer, Fuligni, Lieberman, & Galvan, 2014). One possibility is that increased emotional reactivity in mid-adolescence is associated with heightened reward valuing in prosocial contexts (Telzer, 2016).

Indeed, brain imaging research has allowed for empirical evaluations of prosocial developmental processes by relating neural activity to prosocial behaviors, resulting in three important findings. The first finding pertains to the ventral striatum, a region involved in many different types of reward and learning signals. This region is especially well known for its role in processing a variety of basic rewards (Haber & Knutson, 2010). The ventral striatum has anatomical and functional connections to the orbitofrontal cortex, also referred to as ventral medial prefrontal cortex (Lieberman et al., 2019), and together this network of brain regions has been interpreted as having a crucial role in updating reward values (Delgado et al., 2016). Studies in adults have reported that ventral striatum activity does not only correspond with monetary rewards, but also with feelings of inclusion (Tamir & Hughes, 2018), cooperation (Rilling & Sanfey, 2011), and fairness, suggesting that the ventral striatum is also sensitive to social rewards

(Tabibnia, Satpute, & Lieberman, 2008). This social reward sensitivity is therefore thought to also underlie prosocial motivations, as it reinforces behaviors that benefit others and strengthens social relationships (Fett, Gromann, Giampietro, Shergill, & Krabbendam, 2014). Second, based on research in adults, there is converging evidence from functional neuroimaging studies for a crucial role of the medial PFC, TPJ, and STS (also referred to as the "social brain" network) in situations that require individuals to consider about thoughts and intentions of others, such as helping and trusting others (Amodio & Frith, 2006; Lieberman et al., 2019; Rilling & Sanfey, 2011). These forms of perspective-taking play an important role in motivations that underlie prosocial actions (Crone & Fuligni, 2020). Third, various studies have demonstrated that the dorsolateral prefrontal cortex (DLPFC) plays an important role in goal-flexibility related to balancing between the needs of self and others, for example by inhibiting selfish impulses (Achterberg, van Duijvenvoorde, Bakermans-Kranenburg, & Crone, 2016) or by engaging strategic actions (van den Bos, van Dijk, Westenberg, Rombouts, & Crone, 2009). Together, these processes, which rely on brain regions that develop during adolescence (Blakemore & Mills, 2014), are thought to work in concert when acting prosocially toward various targets and in various contexts.

Two additional processes related to prosocial behavior that are sensitive to individual differences in adolescence are empathy and norm processing. Empathy refers to the communication of an emotional state from one individual to another and is associated with multiple cooperating brain regions, from mirror neurons to cognitive control (Decety & Holvoet, 2021). A recent literature review outlines evidence that empathy in its basic form develops in childhood through interactions with the environment, with an important role for the family context (Brownell, Svetlova, Anderson, Nichols, & Drummond, 2013). These basic empathic abilities are important building blocks for more complex socio-cognitive processes such as perspective taking, which develop further in adolescence (Decety & Holvoet, 2021; Van der Graaff, Carlo, Crocetti, Koot, & Branje, 2018). Norm processing develops considerably in childhood years, with strong social equity norms around the ages of 8–9-years (Meuwese, Crone, de Rooij, & Guroglu, 2015). In adolescence, these equity norms become replaced by more complex norms that require higher levels of perspective-taking (understanding intentions or others) and goal flexibility (taking into account the broader social context) (Guroglu, van den Bos, & Crone, 2014; van den Bos et al., 2010). Prior studies in adults revealed a unique set of brain regions that are responsive to norm violations specifically; the anterior cingulate cortex (ACC) and the bilateral insula (Rilling & Sanfey, 2011). These regions are typically engaged when individuals perform actions that go against their personal norms (Guroglu, van den Bos, van Dijk, Rombouts, & Crone, 2011). Even though empathic concern and norm processes appear to be relatively stable in adolescence in terms of developmental processes, empirical studies typically show large individual differences (Meuwese et al., 2015; Stern & Cassidy, 2018; Van der Graaff et al., 2018), possibly suggesting relatively larger susceptibility to the environment, which Is why we take these processes into account in our model of prosocial development (Decety & Holvoet, 2021).

In the next sections, we implement this model in three steps: (i) we evaluate a multiple-pathway neuroscientific model of prosocial development by relating neuroscience discoveries to developmental changes in key dimensions of prosocial development and their sensitivity to various contextual factors; (ii) we test the role of environmental support factors by reviewing experimentally controlled and naturalistic service learning programs aimed at fostering socio-affective and/or socio-cognitive processes that contribute to prosocial development in adolescence. This approach will allow us to evaluate the important question: (iii) When and how do changes in socio-affective and socio-cognitive development result in opportunities for prosocial development and which factors facilitate opportunities for positive, prosocial development?

4. Developmental neural pathways of prosocial behavior

Although originally regarded as a generalized construct, recent studies have elucidated that prosocial behavior is an umbrella term consisting of many different types of other-benefitting behaviors. These studies have shown that different types of prosocial behaviors do not always correlate within individuals, and often have unique antecedents and developmental patterns (Carlo & Padilla-Walker, 2020; Padilla-Walker et al., 2018). In the next sections, we examine developmental changes in four key dimensions of prosocial development which are increasing in complexity: (i) socio-affective valuing of rewards for others through vicarious gains and cooperation, (ii) socio-cognitive understanding of needs when helping, (iii) combining socio-affective and socio-cognitive building blocks during giving and sharing, and (iv) understanding long-term consequences for others when trusting or reciprocating trust.

The various forms of prosocial behavior are also situated within a multitude of contexts, including variations in the target/recipient of prosocial behavior (e.g., family, friends, community), the visibility of prosocial behavior (e.g., being observed online or by an audience), the strategic context (power of the recipient), and the needs and time periods of prosocial behavior (e.g., targets in need, COVID-19). In the subsequent section, we describe studies that aim to decompose contextual influences on prosocial behavior, which allows us to examine these various domains in more detail (Luo, 2018). These processes will be examined by reviewing behavior and fMRI studies including children, adolescents and adults.

4.1 Valuing rewards for others through vicarious gains and cooperation

One motivation for prosocial actions can be the pleasure of receiving rewards for others (Harbaugh, Mayr, & Burghart, 2007; Morelli, Knutson, & Zaki, 2018). A way to operationalize this is through vicarious rewards, that is, rewards that are received for another individual, either in a mutual gaining context or gaining only for others. Typically, these rewards are non-costly as they do not come at the expense of self.

It is well known that the ventral striatum is a reward center in the brain that responds strongly to receiving rewards for self (Haber & Knutson, 2010). A significant number of studies has found that, relative to children and adults, activity in ventral striatum is heightened during adolescence when receiving rewards for self, suggesting more emotional reactivity in response to reward (Casey, 2015; Galvan, 2010). This result has been replicated several times using a variety of gambling paradigms, such as passive gambling tasks (Galvan et al., 2006; Van Leijenhorst et al., 2010), active gambling tasks (Van Leijenhorst et al., 2010), social risk taking tasks (Chein, Albert, O'Brien, Uckert, & Steinberg, 2011), and probabilistic learning tasks (Cohen et al., 2010), and was confirmed in a meta-analysis (Silverman, Jedd, & Luciana, 2015). Although prior studies have mostly relied on monetary rewards to elicit striatal reactivity, recent work shifted to other forms of rewards, showing that the ventral striatum response appears to be highly sensitive to social factors, especially in adolescence (Chein et al., 2011). One hypothesis is that emotional reactivity in adolescence in terms of ventral striatum activity to vicarious rewards can account for changes in the emotional valuing of prosocial activities (Telzer, 2016; Telzer, Fuligni, Lieberman, & Galvan, 2013). Indeed, prior research in adults showed that the ventral striatum is most responsive to rewards for close others relative to distant others (Morelli et al., 2018), and when there is higher social identification with the group (Hackel, Zaki, & Van Bavel, 2017). Together, these studies show that in adults the ventral striatum may be an important marker for the "warm glow" of receiving rewards for others (Harbaugh et al., 2007; Rilling & Sanfey, 2011). Below, we summarize studies that have examined whether adolescence is a time of heightened reward activity, not only for self but also for others.

In studies focusing on adolescence, vicarious rewards have been examined for family members, friends, unknown others such as peers, and broader community partners (e.g., charity). First, in a cross-sectional study vicarious neural reward responses were examined for family members, specifically mothers. Adolescents aged 10–27-years gained money for self or for their mother, and a neural peak ventral striatum in mid-adolescence was observed when vicariously gaining for mothers (Braams & Crone, 2017). Vicarious gains may be an important factor in valuing the outcome of cooperation, which refers to a group of individuals working together toward a similar goal. To examine the relation between neural activity in the ventral striatum for vicarious gains in the family context in more detail, a second study examined vicarious gains in a Prisoner Dilemma format (Brandner, Guroglu, & Crone, 2020). The Prisoner Dilemma Game is a cooperation game where two players each decide simultaneously whether to cooperate or defect. In case of mutual cooperation, both players receive a moderate size reward (Rilling et al., 2002). An experimental behavioral part of the study in 9–18-year-old adolescents revealed differential developmental trajectories for cooperation with parents (increasing with age) and unknown others (peaking in mid-adolescence followed by a decrease in adulthood) (Brandner, Guroglu, van de Groep, Spaans, & Crone, 2021; see also Box 1). A false-choice fMRI version of the Prisoner Dilemma Game administered to the same participants showed that ventral striatum activity scaled with reward values for self, but the ventral striatum was also responsive to vicarious rewards for parents (see Fig. 2A; Box 1). In contrast, no such vicarious reward response was observed when gaining for unknown peers (Brandner et al., 2020, 2021). Together, these findings show evidence for vicarious neural gains for family members, with some evidence that this activity is heightened in mid-adolescence (Braams & Crone, 2017) and scales with pleasure of winning for mothers (Brandner et al., 2021).

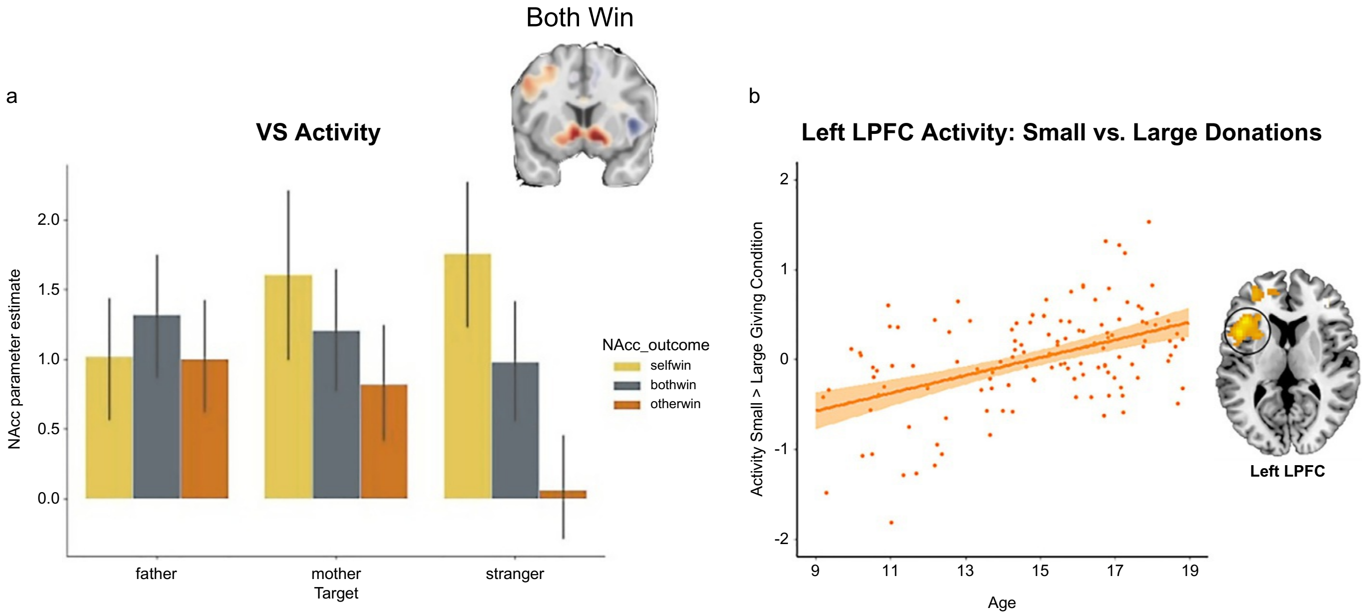


Fig. 2 Example of neural activity in a social-affective vicarious reward task (Brandner et al., 2021) and socio-cognitive giving task (van de Groep et al., 2022) based on the first wave of the Brainlinks study (Box 1). (A) Results showing heightened ventral striatum activity when gaining for family members, but not for strangers in adolescence. (B) Young adolescents engage the dlPFC stronger for large donations whereas older adolescents engage dlFPC more for small donations.

Second, a longitudinal study examined whether vicarious gains for friends elicited a similar developmental pattern of neural activity as rewards for self. For this study, 8–29-year-old participants reported who was their best friend on three longitudinal waves separated by 2 years. This allowed for the distinction between adolescents with stable best friends (same friend across 3 waves) and unstable best friends (different best friend across 3 waves). For adolescents with stable best friends, there was a peak in neural activity in mid-adolescence for vicarious rewards for friends as well as for rewards for self. In contrast, no such developmental pattern was observed for adolescents with unstable best friends, although for unstable best friends, ventral striatum activity correlated with experienced friendship quality (Schreuders, Braams, Crone, & Guroglu, 2021). These findings show evidence for vicarious neural gains for close friends in adolescence.

Third, vicarious rewards can be gained also for more distant prosocial partners with whom the participant does not have a direct connection. One such recipient can be a charity, which typically receives prosocial actions because of the observed need and because charities are considered societal trustworthy recipients (Harbaugh et al., 2007). Using a similar Prisoner Dilemma Game format, one recent study including adolescents aged 11–21-years showed that when gaining vicariously for charity, on the group level charity gains were not associated with increased activity in the ventral striatum (Spaans, Peters, & Crone, 2019). However, it was found that adults who scored higher on self-reported empathy (Spaans, Peters, & Crone, 2019) and that adolescents who scored higher on perspective taking and donation behavior (Spaans, Peters, & Crone, 2020), showed higher activity in the ventral striatum when vicarious gaining for charity, possibly suggesting that they feel a closer connection to the charity. Together, these findings show that ventral striatum activity is related to the relationship with the recipient, with higher activity when the target is experienced as closer or when the recipient is deserving, such as in the case of charity.

Finally, to examine whether vicarious rewards were related to behavioral adaptations, prior studies examined vicarious gains in a learning paradigm. In one functional neuroimaging study, it was examined whether behavioral learning rates and neural prediction errors in a probabilistic learning task for self and unknown peers differed across 9–21-year-old adolescents (Westhoff, Blankenstein, Schreuders, Crone, & van Duijvenvoorde, 2021). A prior study in adults already showed that prediction errors for unknown peers were related to activity in the ventral striatum, but only activity in the subgenual ACC, a region bridging the ventral striatum and the vmPFC, was correlated to self-reported empathy (Lockwood, Apps, Valton, Viding, & Roiser, 2016). In a similar study including adolescents and adults, it was found that learning rates were higher in younger adolescents (indicating more immediate adaptation) when learning for peers. At the neural level, ventral striatum activity was higher for prediction errors for self than for peers across adolescence, but activity in vmPFC showed an age-related increase when learning for others (Westhoff et al., 2021). An intriguing question for future research is therefore to examine how vicarious gains and learning rates develop for close others, such as family and friends.

Taken together, in a vicarious reward setting, adolescents show neural peaks in activity when receiving rewards for close others such as mothers and stable friends, but not for more distant others such as charity, unstable best friends, or unknown peers. In case of more distant others, ventral striatum activity correlates more strongly with perspective taking (charity), friendship quality (unstable friends), and prediction errors (unknown others).

4.2 Helping: Social-cognitive perspective taking

When prosocial behaviors involve an action to contribute to the needs of others, this can be defined as helping. In its simplest format, helping behavior is non-costly as it does not need to involve giving up resources to provide assistance to others. Helping does, however, involve an understanding of the needs of others (Carlo & Padilla-Walker, 2020; Decety & Holvoet, 2021).

Prior studies have examined helping behavior in the context of needs of unknown peers in a prosocial Cyberball Game. The traditional Cyberball game involves a three player ball tossing game where one of the players is excluded, leading to negative feelings and loss of control (Boyes & French, 2009). In the prosocial Cyberball game, there are four players where the participant observes that two players exclude a third player from the ball tossing game. The participant has the possibility to help the excluded player by increasing the number of ball tosses toward them, thereby compensating for their exclusion (Riem, Bakermans-Kranenburg, Huffmeijer, & van IJzendoorn, 2013). In a behavioral study, it was found that adolescents between ages 9–17-years compensate for exclusion, but compensation was higher for adolescents who reported more empathy (Vrijhof et al., 2016). In an fMRI study including adults, it was observed that tossing to the excluded player was associated with increased activity in the ventral striatum and temporal parietal junction (van der Meulen, van IJzendoorn, & Crone, 2016). A similar set of studies in children, in contrast, showed either no robust neural activity in three studies including children ages 7–10-years (van der Meulen et al., 2017) or increased activity in the precuneus when compensating exclusion in 7–8-year-old (van der Meulen, Steinbeis, Achterberg, van IJzendoorn, & Crone, 2018). A different study comparing 12–17-year-old adolescents with 22–30-year-old adults showed that adults more often helped excluded others than adolescents, which was accompanied by more activation in the TPJ and medial prefrontal cortex in adults (Tousignant, Eugene, Sirois, & Jackson, 2018).

Together, these studies suggest that helping excluded unknown peers is associated with activation in different neural regions depending on the age of the participants, specifically showing increased activity in the TPJ in adults relative to children and adolescents (Tousignant et al., 2018; van der Meulen et al., 2016). Even though it has not yet been examined how friends are compensated during adolescence, a prior behavioral study showed that participants only help unknown peers when they are being excluded by unknown others, but not when they are excluded by their friends (Spaans, Will, van Hoorn, & Guroglu, 2019). Thus, the extent to which adolescents show helping behavior depends on the development of perspective taking and the context in which help is needed, such as whether the excluded target is a family member, friend, or unknown other.

4.3 Giving: Socio-affective and socio-cognitive building blocks

Giving is a costly prosocial act in which an individual distributes valuable resources between themself and someone else (Cutler & Campbell-Meiklejohn, 2019). There is accumulating evidence from studies in adults that this behavior is driven by both socio-affective and socio-cognitive processes, which are represented in intuitive and deliberative neural systems, respectively (Feng, Luo, & Krueger, 2015; Luo, 2018). The intuitive, socio-affective system includes regions such as the ventral striatum, vmPFC, and anterior insula, which play a role in the processing of reward valuing and norm violations, respectively (Luo, 2018). The more deliberate socio-cognitive system includes regions such as the dLPFC, TPJ, and STS) (Cutler & Campbell-Meiklejohn, 2019; Feng et al., 2015; Luo, 2018). Below, we give an overview on the role of these affective and cognitive brain systems in giving behavior, which often requires adolescents to take the target of giving and situational demands into perspective.

One way to operationalize giving behaviors is by utilizing economic games, which are structured experiments which model interdependent situations (Cutler & Campbell-Meiklejohn, 2019; Gummerum, Hanoch, & Keller, 2008) (see Fig. 3). The Dictator Game is the most basic economic game, in which one individual decides upon a certain split of valuable resources (Engel, 2011). The target or recipient of this division has no power over this decision, hence the name Dictator Game. Although economic games were initially developed with the idea that individuals would show self-interested, rational (i.e., homo-economical) behavior, studies in adults and adolescents alike quickly showed that individuals tend to give away 20–30% of their resources to unknown others, even in situations where there are no extrinsic or future rewards associated with this decision (Urbina & Ruiz-Villaverde, 2019). Neuroimaging studies have often utilized variations of the Dictator Game to study giving behavior as its structured nature enables the quantification of this complex behavior. This allows for the comparison of neural activity associated with varying levels of generosity. These studies, which have mostly been performed in adults, have elucidated that, compared to selfish decisions, giving elicits activation in various regions, including the nucleus accumbens, mPFC, and dlPFC (Cutler & Campbell-Meiklejohn, 2019).

Giving in the Dictator Game: One study in 8–16 year-old children and adolescents operationalized giving by comparing four conditions: costly giving, non-costly giving, costly-rewards, and non-costly rewards to unknown peers (Do, McCormick, & Telzer, 2019). Comparing costly giving to costly- and non-costly rewards revealed no differences in neural activation. In contrast, comparing costly to non-costly giving revealed activation in the precuneus and inferior temporal gyrus. A subsequent analysis examined costly prosocial versus non-costly prosocial choices in a subsample of youth that made sufficient prosocial decisions. This analysis revealed a quadratic peak in neural activation in the pSTS, temporal pole and inferior frontal gyrus (IFG) in early adolescence when comparing costly giving to costly rewards; and in pSTS, dlPFC, and IFG activation when comparing costly versus non-costly giving. Overall, these results suggest elevated activation in socio-cognitive systems of the brain in early/mid-adolescence, particularly in relatively prosocial individuals, possibly reflecting higher goal-flexibility.

In a recent study from our own group (van de Groep et al., 2022; van de Groep, Zanolie, & Crone, 2020a; for more details see Box 1), we designed a variation of the Dictator Game in which individuals could give away coins in either a small or large giving condition, to control for the number of

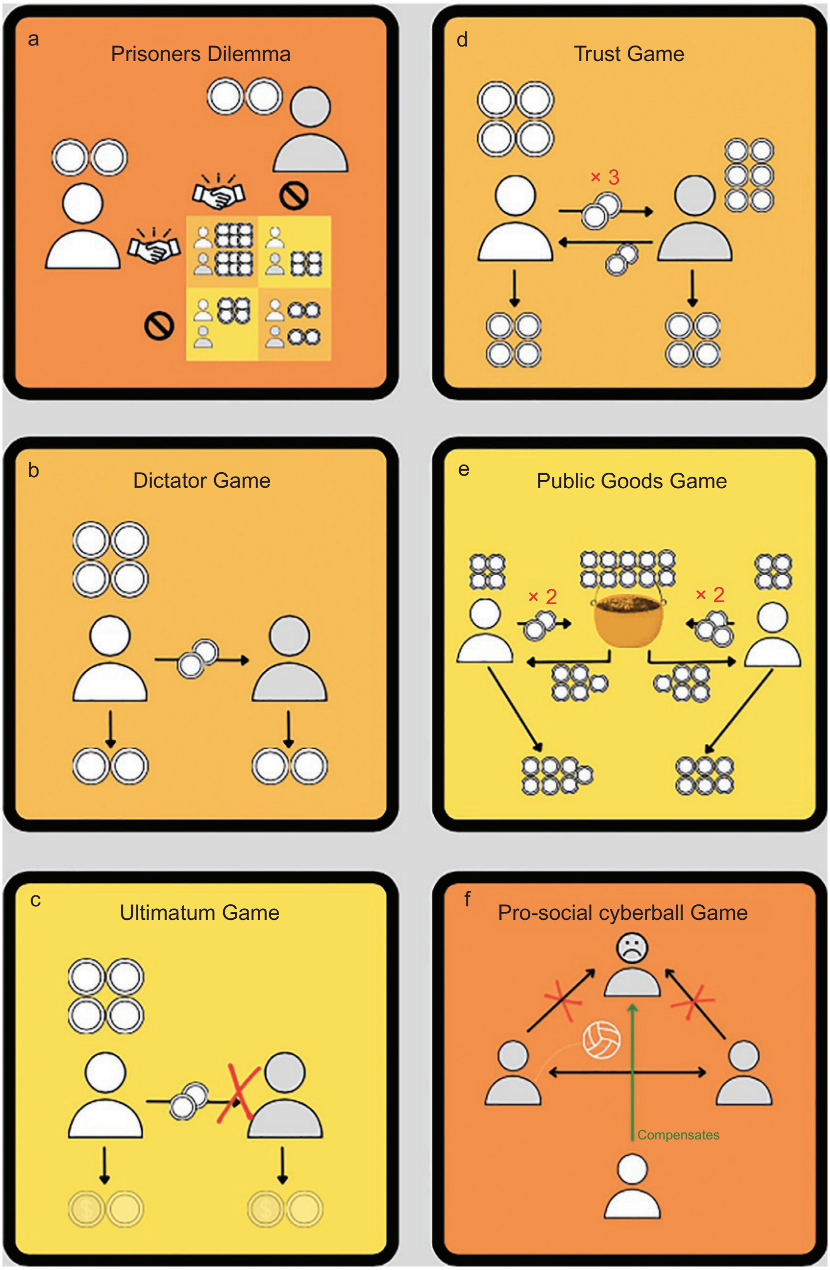


Fig. 3 See figure legend on opposite page.

prosocial choices. In the small giving condition, individuals could give away 1, 2, or 3 out of 7 coins. In the large giving condition, individuals could give away 4, 5 or 6 out of 7 coins. This design allowed us to compare small versus large size giving conditions, but also relative generosity within these giving conditions. Importantly, this design also allowed for voluntary decisions within the small and large giving conditions, as voluntary decisions give the best indication of generosity (Gagné, 2003). The results of the first wave of this longitudinal study ($N = 128$, ages 9–19) revealed that adolescents gave more in the small compared to large size giving condition. Giving very small or very large amounts was associated with increased activity in the mPFC and anterior insula, suggesting a general response to giving size. In addition, age comparisons revealed that older adolescents showed increased lateral and anterior PFC activation for small size giving (van de Groep et al., 2022) (see Fig. 2B; Box 1). These results were interpreted to suggest a role for the mPFC and anterior insula in saliency detection and norm processing, and show a developmental shift from stronger activity in the anterior and lateral PFC for large (high-costly) giving to small (low-costly) giving. These studies fit well with recent studies comparing Ultimatum Game giving (i.e., strategic) with Dictator Game giving. The Ultimatum Game is a game with the same structure as the Dictator Game but where the second player has the possibility to reject the offer, in which case both players receive nothing. Comparing this additional strategic element of the Ultimatum Game versus the Dictator Game resulted in elevated activity

Fig. 3 Example of economic games that are often used to examine prosocial behavior. (A) The Prisoner Dilemma Game is a cooperation game where a mutual cooperation choice leads to mutual benefit. (B) The Dictator Game is an economic game where a player can divide resources between themselves and a second player. (C) The Ultimatum Game is an economic game that is similar to the Dictator Game, but with the possibility of the second player to reject in which case both players receive nothing. (D) The Trust Game is a two-player exchange game, where the first player can decide on the division of resources, or can trust the second player in which case resources are increased. The second player has the option to reciprocate (the prosocial choice) in which case both players moderately benefit, or to defect in which case the second player benefits most. (E) The Public Goods Game is a multi-player economic game where all players contribute resources to a common good, which is then increased by a multiplier and shared among the players. The prosocial game is to contribute. (F) The Prosocial Cyberball Game is a helping paradigm where the participant has the possibility to compensate (i.e., help) a player who is being excluded by other players in a ball tossing game.

in lateral PFC for strategic giving, which increases between ages 6–12-years (Steinbeis, Bernhardt, & Singer, 2012, see also Guroglu et al., 2011).

Taken together, costly and strategic giving involve both socio-affective and socio-cognitive neural systems in deciding upon giving size (i.e., generosity). These studies show possible evidence for higher goal flexibility as demonstrated by increased lateral PFC activity in early to mid-adolescence (Do et al., 2019; van de Groep et al., 2022).

Giving to various targets: Giving to unknown others tells only part of the story of human generosity, as most of our interactions are with people we know (Guroglu et al., 2014). Several studies have focused on relational giving, such as giving to family and friends.

One study showed that neural activation related to giving to one's family depended on their culture (Telzer, Masten, Berkman, Lieberman, & Fuligni, 2010). Whereas Latino and White participants showed similar levels of generosity toward their family, they showed distinct patterns of activity within the mesolimbic reward system. Latino participants showed more reward activity when contributing to their family, while White participants showed more reward activity when they gained money for themselves. Reward activity was measured as activity in the ventral and dorsal striatum, as well as the ventral tegmental area. These results show that characteristics of the benefactor and target of giving interact to shape giving decisions.

Other studies have examined giving to peers, such as a recent study that examined giving to real-life friends, disliked peers, neutral peers, and unfamiliar peers in mid-adolescence (Schreuders, Klapwijk, Will, & Guroglu, 2018). Here, adolescents showed highest levels of giving to friends, and lowest levels for disliked peers. Giving to friends compared to disliked peers was associated with activation in the putamen and posterior middle temporal gyrus, and giving to friends compared to unfamiliar peers was associated with activation in the superior parietal lobule and precentral gyrus. However, these studies did not examine age-related differences in neural activation in relational giving.

In our own study we aimed to examine age-related differences in neural activation pertaining to giving to friends and unfamiliar peers (van de Groep et al., 2022, see also Box 1). To this end, we employed the aforementioned variation of the Dictator Game in which participants could divide seven coins between friends and unfamiliar peers, in either a small or large giving condition. Here, we observed that individuals gave more to friends than to unfamiliar peers. This differentiation was greater in older compared to younger adolescents. Giving to friends was associated with activation in the right insula, bilateral TPJ, right lateral PFC, and right SMA, but no age-related differences were observed in neural activity for friends compared to unfamiliar peers. One region, the precuneus, showed an interaction between giving size and target, such that activation was lowest for large-size giving to unfamiliar peers, suggesting that adolescents were less likely to engage in perspective taking during these decisions compared to decisions that involved small-size giving and giving to friends. Overall, these results suggest that age-related differences in giving in adolescence may be specific to friends, a target whose importance greatly increases during this period in life (Van Hoorn, Van Dijk, Guroglu, & Crone, 2016). However, it remains an empirical question whether and how this bias in giving to friends is reflected in the brain. Answering this question may require a different methodology, such as longitudinal analyses or functional connectivity analyses.

Finally, neuroimaging studies have investigated the extent to which neural activity is modulated by in-group to out-group giving. These studies in young adults (Telzer, Ichien, & Qu, 2015) and adolescents (Do & Telzer, 2019) have shown that youth give more to in-group members (i.e., have an in-group bias). In young adults, this was accompanied by increased activation in the ventral striatum. Moreover, a greater sense of group identity was associated with heightened activation in the VLPFC, ACC, TPJ, and dmPFC when contributing to out-group compared to in-group members, suggesting that this requires additional self-control and perspective taking related processing (Telzer et al., 2015). The 8–16-year-old adolescents were more likely to give when there was a greater discrepancy between outcomes for others over oneself (i.e., higher reward inequity). No differences in general corticostriatal activation were observed for giving to in-versus out-group members, but adolescents showed greater connectivity between the ventral striatum and posterior STS when considering relatively inequitable decisions that benefited out-group peers (Do & Telzer, 2019). This study observed no age differences, suggesting that in-group versus out-group biases and the associated brain processes already exist from childhood onwards.

Giving in situational context: Giving can require individuals to take into account other social situational demands in addition to the target. One example is whether decisions are being observed by a peer audience. One study in adolescents, in which 12–16-year-olds played a public goods game (i.e., they divided valuable resources between themselves and a group) shows that adolescents gave more to the group (at their own expense) when they were being observed, and even more when they received evaluative feedback from peers (Van Hoorn, Van Dijk, Guroglu, & Crone, 2016). Peer presence was associated with activation in the mPFC, TPJ, precuneus, and STS, suggesting socio-cognitive and perspective taking related processing. In this study, younger adolescents were more generous, but peer presence effects did not differ as a function of age. Finally, adolescents' TPJ activity was associated with generosity. This is in line with another study which observed that individual differences in TPJ recruitment while viewing others' prosocial behaviors were associated with adolescents' own charitable giving (Tashjian, Weissman, Guyer, & Galvan, 2018), suggesting an important role for the TPJ in balancing the needs of self and others. In our own work, we also examined Dictator Game giving decisions for friends and unfamiliar peers in audience and anonymous contexts: in half of the trials, decisions would be completely anonymous, whereas in the other half, decisions would be observed by peers later in time (van de Groep et al., 2022; see also Box 1). Adolescents were more generous in the audience condition, and the difference in activation in the insula for friends compared to unfamiliar others was amplified in the audience but not anonymous condition, suggesting that peer presence can increase social concern or saliency for friends.

All in all, the current literature on giving shows that giving in adolescence is highly dependent on the social context, including the target and situation. Studies that examined relative generosity and social contextual factors, such as target and peer presence, reveal that giving is associated with neural activation in regions involved in both socio-affective (i.e., ventral striatum, mPFC, and insula) and socio-cognitive processes (i.e., TPJ, lateral PFC, STS). This suggests that giving decisions are shaped by balancing feelings (e.g., related to reward or saliency) and cognition (e.g., perspective taking, goal flexibility) associated with outcomes for self and others (Crone & Fuligni, 2020). There is evidence that adolescents recruit these regions differently than adults. Sometimes, adolescents show reduced activity in the TPJ when performing a prosocial task, for example when helping (Tousignant et al., 2018). However, in other situations adolescents show elevated activity in the lateral PFC, for example when giving (Do et al., 2019). We hypothesize that these differences reflect higher goal flexibility in adolescents relative to adults, with flexible recruitment of PFC, TPJ and STS, depending on whether situations require exploration, perspective taking, and forming new social connections (Casey, 2015; Crone & Dahl, 2012).

4.4 Trust/reciprocity: Contribution of multiple processes

Trust and reciprocity are two important, potentially costly, prosocial processes that are more complex than giving behavior, as they require an interaction between a trustee and trustor (Rilling & Sanfey, 2011). As a result, trust and reciprocity require advanced and mature levels of perspective taking and strategic thinking, processes that still develop in adolescence and young adulthood. Trust, defined as decisions favoring outcomes for other individuals aiming at future cooperation and self-gain, is important for the development of positive relationships, whereas reciprocity, defined as a mutual exchange (e.g., reciprocating trust shown by another individual), is important for maintaining these relationships (Lahno, 1995). Trust and reciprocity behavior are therefore types of prosocial behaviors that are more oriented toward long term relationships. Developing trust and reciprocity behavior helps adolescents to successfully navigate a complex social world, which aids them to develop new social relations based on values such as cooperation and sharing.

An economic game that is often used to study trust and reciprocity is the Trust Game (Berg, Dickhaut, & McCabe, 1995). In the Trust Game, two players are involved in dividing a certain number of resources. The first player (trustor) chooses how to divide the initial number of resources between themselves and the second player (trustee). The given number of resources by the trustor is multiplied and given to the trustee. The trustee then decides how much they want to give back to the trustor. Studies examining trust and reciprocity using the Trust Game have shown age-related increases between childhood and adolescence in trust decisions (van den Bos et al., 2010). However, other studies have shown a general stability in trust and a decrease in reciprocity (van de Groep, Meuwese, Zanolie, Guroglu, & Crone, 2018). The exact developmental patterns seem highly dependent on the level of perspective taking required (van de Groep et al., 2018; van den Bos et al., 2010), and, in case of iterative games, on the behavior of the other players (Fett, Gromann, et al., 2014; Westhoff, Molleman, Viding, van den Bos, & van Duijvenvoorde, 2020).

Trust and reciprocity behavior are more strategic than other types of prosocial behaviors, because of the second player involved in the social context. Deciding whether to show trust and reciprocity toward another individual is dependent on socio-cognitive processes, such as perspective taking, risk calculation, and outcome monitoring (Burke, van de Groep, Brandner, & Crone, 2020). Just like giving, trust and reciprocity decisions involve both socio-affective and socio-cognitive processes. However, given the additional socio-cognitive demands of such a strategic task, it was found that an especially important role is reserved for the more socio-cognitive developmental processes, including underlying brain networks such as the dlPFC, TPJ and STS. Indeed, a prior study demonstrated an age-related increase in dlPFC when receiving trust, as well as an increase in TPJ during reciprocity, which requires perspective taking (van den Bos, van Dijk, Westenberg, Rombouts, & Crone, 2011).

Although most studies using the Trust Game have focused on trust and reciprocity toward unknown others, recent evidence demonstrates that these behaviors depend on the target. A recent study using an iterative Trust Game with partners of varying levels of trustworthiness demonstrated that adolescents show higher levels of trust and reciprocity choices when the other individual is more trustworthy (i.e., in a cooperative context) relative to a less trustworthy interaction partner (Fett, Gromann, et al., 2014). These findings illustrate that adolescents learn over time whom to trust, which was associated with increased activity in the TPJ and the precuneus, regions previously associated with perspective taking (Carter & Huettel, 2013). In addition, it has been observed that adolescents show higher levels of trust and reciprocity toward friends compared to unknown, neutral, and disliked peers (Guroglu et al., 2014).

Whereas previous studies mainly focused on trust and reciprocity choices toward targets that differentiated in closeness and trustworthiness, such as family, friends, and unknown others (Guroglu et al., 2014), currently less is known about this prosocial behavior oriented toward more distant others and society. Adolescence is an important developmental period for expanding the social world and acquiring societal values (Crone & Fuligni, 2020). Given that this also touches upon adolescents' fundamental need to contribute (Fuligni, 2019), the development of trust and reciprocity behavior toward society may be formative for adolescents (Fuligni, 2020). Future studies may therefore investigate how the social context, particularly who the other is, moderates the developmental pattern of trust and reciprocity behavior during adolescence. Another Important future direction Is to study the role of flexible goal recruitment, as reflected by flexible PFC, TPJ and STS recruitment, when considering whether to engage in prosocial behavior such as showing trust and reciprocity to others (see Box 1).

5. Environmental influences on prosocial behavior

Given the separable developmental time courses of socio-cognitive and socio-affective processes, and described differences in antecedents of prosocial behaviors, it may be expected that during adolescence, some aspects of prosocial behavior are specifically sensitive to environmental influences. We recently showed that, even though prosocial behaviors are relatively stable across time in early and late adolescence, socio-affective (i.e., empathy) and socio-cognitive (i.e., perspective taking) processes are differentially related to prosocial behaviors such as giving, altruism, and emotional support (te Brinke, van de Groep, van der Cruijsen, & Crone, n.d.). This leads to the question whether individuals are more sensitive to environmental influences in adolescence, during periods of higher social-affective reactivity and goal flexibility. In this section, we first present intervention findings, examining whether prosocial experiences are important for wellbeing, contribution to society, and forming meaningful relationships with others. In the second part of this section, we examine how prosocial behaviors are shaped by family, peer, and societal contexts.

5.1 Intervention effects

Several studies have examined whether prosocial behavior can be fostered during adolescence using intervention designs. Meta analyses show that, overall, these interventions have positive effects on both participating adolescents and their environments (Curry et al., 2018; Mesurado, Guerra, Richaud, & Rodrigues, 2019; Shin & Lee, 2021). Studies that aim to promote prosocial behavior in children, adolescents and/or adults by instructing them to perform "Acts of Kindness" (e.g., greet unknown peers in the hallway, spend money on someone else) show a small to medium positive effect of these prosocial actions on the wellbeing of the actor (Curry et al., 2018). Moreover, intervention programs that aim to improve the prosocial behavior of children and adolescents, have a moderate positive effect on prosocial behavior outcomes (Mesurado, Guerra, et al., 2019). Finally, a meta-analysis examining the effects of prosocial behavior interventions on adolescents, also found a small to moderate positive effect on prosocial behavior outcomes (Shin & Lee, 2021). Intervention effects do not appear to be moderated by age or grade levels (Curry et al., 2018;

Mesurado, Guerra, et al., 2019). There are, however, large differences in the focus, duration, and targets of these interventions. In the next part, we will therefore discuss these programs.

A first difference between intervention programs is their relative focus on socio-affective and/or socio-cognitive skills. Interventions that aim to enhance socio-affective skills such as empathy, frequently use a combination of experience- and practice-based learning skills (i.e., experiencing empathy through videos or real-life experiences, engaging in prosocial activities). An example of an intervention program that focuses on socio-affective skills is the "Roots of Empathy" program (Schonert-Reichl, Smith, Zaidman-Zait, & Hertzman, 2012). This 9-month program is implemented by elementary school teachers and includes a monthly classroom visit by an infant and his/her parent(s) whom the class "adopts" at the beginning of the school year. Research shows that this intervention has a positive effect on the peer nominated prosocial behavior of 8–12 year-old children. No direct effects on empathy were observed (Schonert-Reichl et al., 2012). Another socio-affective skills intervention is the "Caring for Life" curriculum, which is found to promote teacher-reported prosocial behavior among Chinese elementary school students (Samuels, 2018). In contrast, the intervention program "Try volunteering"—an 8-week program that stimulates adolescents to start volunteer work—has a positive effect on empathy among 13–16 year-olds (Truskauskaitė-Kunevičienė, 2016). The self-administered online "Hero" intervention that stimulates socio-affective skills, such as empathy, emotion recognition, and forgiveness, is found to be effective in promoting prosocial behavior toward family members and unknown others, but direct effects on the targeted socio-affective skills were not assessed (Mesurado, Distefano, Robiolo, & Richaud, 2019). Lastly, an 11-day online prosocial intervention that included motivational videos and daily prosocial behavior exercises was found to have positive effects on both empathy and prosocial behavior among 16–25 year-olds (Baumsteiger, 2019).

In contrast, other interventions have a stronger focus on socio-cognitive skills such as perspective taking. These interventions frequently include skill-based learning strategies, such as exercises that aim to increase adolescents' recognition of the needs and perspectives of others. For example, CEPIDIA, an Italian prosocial behavior intervention, includes five domains of skill-based learning strategies: perspective taking, prosocial values, emotion-regulation, interpersonal-communication, and civic-engagement (Caprara et al., 2014). During early adolescence, this program is found to be effective in promoting helping behavior toward friends and unknown others, both directly after the intervention and at an 18-month follow-up (Caprara et al., 2014; Caprara, Luengo Kanacri, Zuffianò, Gerbino, & Pastorelli, 2015). Moreover, an intervention that targets perspective taking and empathy (and thereby socio-cognitive and socio-affective processes), was found to have positive effect on both perspective taking capacities and prosocial behavior in a sample of 10–12 year-old children (Etxebarria et al., 1994).

A second difference between intervention programs is their relative focus on peer- or adult-based delivery of the intervention. A peer-based intervention approach stems from the finding that adolescents, in contrast to younger children, frequently resent or reject adult-driven interventions because these interventions do not align with their desire to feel respected and to be accorded status (Yeager, Dahl, & Dweck, 2018; Yeager, Fong, Lee, & Espelage, 2015). Several intervention programs have been developed that aim to enhance prosocial behavior via positive peer network influences (Veenstra & Laninga-Wijnen, 2021). For example, the Meaningful Roles intervention aims to increase opportunities for prosocial behavior by providing adolescents specific school jobs (i.e., pupil responsibilities). These school jobs are embedded in a context of frequent verbal and written recognition for prosocial behavior from peers (Ellis, Volk, Gonzalez, & Embry, 2016). Moreover, a cooperative-play intervention, which stimulates positive peer influence, was found to increase prosocial behavior among 10–11 year-olds (Garaigordobil, 2008). An important question for future research is to examine the relative effectiveness of peer- versus adult-delivered prosocial behavior interventions. It may be expected that peer-led interventions are most effective when they focus on implicit peer influence (i.e., imitation and social norms) and target popular peers as role models (Veenstra & Laninga-Wijnen, 2021).

A third difference is whether interventions focus on intra-personal or intra-societal prosocial behavior. In the experience-, skill-based, and peer-delivered interventions that we described above, the desired outcomes were often an increase in intra-personal prosocial behavior, for example behaviors directed to peers (e.g., helping peers in the classroom; (Caprara et al., 2014)) or unknown others (e.g., volunteering; (Truskauskaitė-Kunevičienė, 2016)). However, research from the civic engagement literature shows that interventions that aim to enhance prosocial behaviors toward larger societal goals, may also yield positive effects. Civic engagement refers to prosocial and political contributions to community and society (Wray-Lake, DeHaan, Shubert, & Ryan, 2019). An example of a civic engagement intervention is the action-oriented "Generation Citizen curriculum," in which students choose a local issue that they wish to tackle, learn strategies and skills to take action, and develop an implementation plan (Ballard, Cohen, & Littenberg-Tobias, 2016; Pope, Stolte, & Cohen, 2011). This program is found to be effective in promoting civic self-efficacy; meaning that after following the program, middle and high school students had stronger beliefs in their ability to make a difference in their community, and believed that their voice would be heard if they would speak up about an issue in their community (Ballard et al., 2016). An interesting direction for future research is to examine the interrelatedness of prosocial development and civic engagement in relation to the formative phase of adolescence.

5.2 Shaping prosocial behaviors by family, peer, and societal contexts

In the previous section, it was demonstrated that interventions can modify prosocial behaviors. Adolescents' prosocial behavior intentions, however, do not develop in isolation. Both family members, peers, and broader societal experiences contribute to these developmental patterns (see Fig. 4B). Research shows that parental volunteering is a strong predictor of adolescent volunteering (Andolina, Jenkins, Zukin, & Keeter, 2003; McGinley, Lipperman-Kreda, Byrnes, & Carlo, 2010; Smetana & Metzger, 2005). Moreover, 12–20 year-old adolescents volunteer more frequently when their family is characterized by strong civic orientation values and open communication (van Goethem, van Hoof, van Aken, Orobio de Castro, & Raaijmakers, 2014), and 13–17 year-old adolescents who perceive higher levels of support from their parents are found to behave more prosocially 1 year later (Malonda, Llorca, Mesurado, Samper, & Mestre, 2019). Family influence on prosocial behavior also appears to continue during late adolescence and early adulthood. For example, a study among 18–25 year-olds showed that parenting practices (i.e., parental support and challenge) were positively related to prosocial behavior toward friends and family members (Mesurado & Richaud, 2018). Thus, family members can have a positive impact on adolescents' prosocial development, both through modeling (i.e., parental volunteering) and socialization (i.e., supportive parenting).

Research on peer influences shows that adolescents are also more likely to set prosocial goals or show prosocial behavior, when their (close) friends also value or engage in these behaviors (Barry & Wentzel, 2006; van Goethem et al., 2014). Moreover, 13–17 year-old adolescents who receive

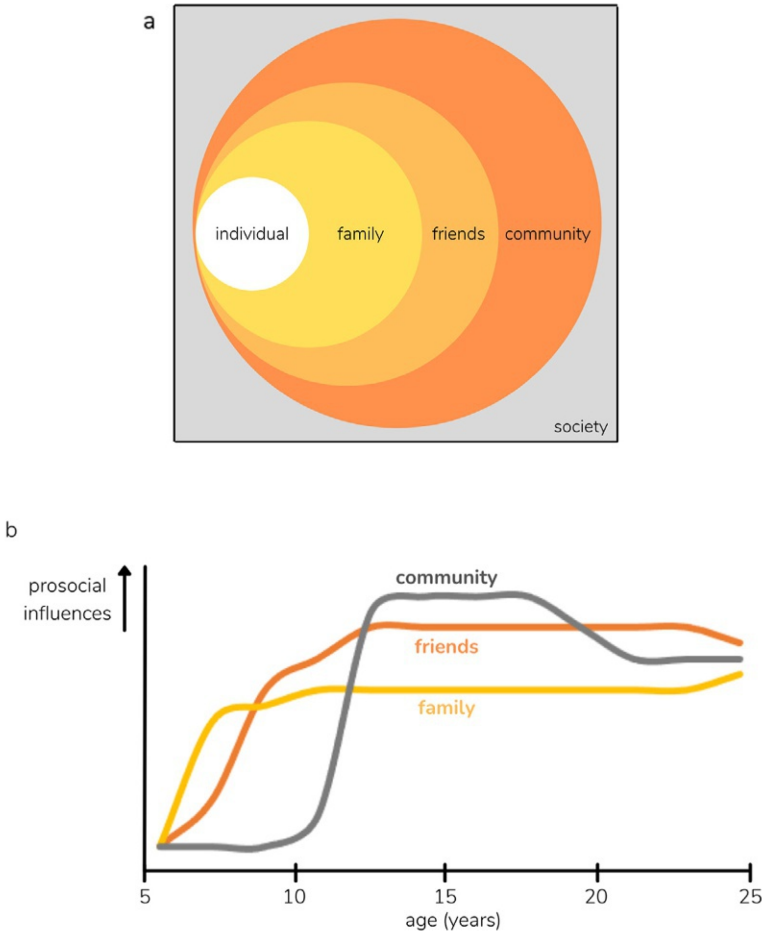


Fig. 4 (A) Example of the various contexts in which adolescents engage and that may shape their prosocial motivation and behavior. (B) Conceptual model showing the developmental pathways for influences on prosocial behavior by family, friends or community targets. Family should be interpreted as influences by parents, and friends should be interpreted as (dyadic) friendships. Community should be interpreted in terms of peer networks, school and neighborhood activities. We expect that influences experiences in one domain shape influences in other domains.

higher levels of support from their friends act more prosocial 1 year later (Malonda et al., 2019). Aside from indirect peer influences (i.e., modeling and socialization), adolescents' prosocial behavior can also be directly influenced by peers. For example, 12–16 year-old adolescents who received manipulated 'prosocial feedback' from friends, are found to give more in a

Public Goods game (Van Hoorn, Van Dijk, Meuwese, Rieffe, & Crone, 2016). A separate study showed that prosocial influence effects are larger in adolescence than in adulthood (Foulkes, Leung, Fuhrmann, Knoll, & Blakemore, 2018), Finally, an experimental study examined the effect of peer influence on prosocial behavior of 12–15 year-olds with an experimental "Chat Room" paradigm. The results showed that adolescents' prosocial intentions were higher after viewing peers behaving prosocially. This effect was moderated by peer status. Specifically, the effects of peer influence were stronger when adolescents interacted with high-status – in comparison to low-status peers (Choukas-Bradley, Giletta, Cohen, & Prinstein, 2015).

There is convincing evidence that family and peer influence shape prosocial behavior. However, prosocial behavior may also be influenced by societal factors (Fig. 4A), such as government or municipal influences, although these are more difficult to study. A recent large societal change was the sudden start of the Covid-19 pandemic. Researchers examined in experimental and survey designs whether these large societal changes affect the behaviors of children, adolescents, and adults. One study showed that 10–12 year-old adolescents who had more face-to-face and virtual societal connections were more willing to help unknown peers during the pandemic (Sabato, Abraham, & Kogut, 2021). In a recent study including adolescents ages 10–25 years, we examined the frequency with which adolescents provided emotional support in the beginning and during the pandemic (Sweijen et al., n.d.). It was found that providing emotional support peaked in mid-adolescence, consistent with prior studies (Blankenstein, Telzer, Do, van Duijvenvoorde, & Crone, 2020). We also observed emotional support increasing during the pandemic, possibly because adolescents created new opportunities to help others after an initial hard lockdown. Finally, the study made use of a giving paradigm, using modifications of the Dictator Game, where the targets were medical doctors (deserving targets), Covid-19 patients, individual with poor immune systems (needing targets), friends, and unknown others. Adolescents gave most to needing and deserving targets, fair splits to friends and least to unknown partners, consistent with a prior study (van de Groep, Zanolie, Green, Sweijen, & Crone, 2020). Giving to needing and serving targets was higher in adolescence compared to adulthood (Sweijen et al., n.d.).

An interesting direction for future research will be to examine how influences by immediate interactions (parents, peers) and larger societal influences (government regulations) influence the trajectories of prosocial development. The Covid-19 crisis is not the only crisis that affects adolescents; they are also confronted with rapidly changing climates and increasing societal inequalities. As stated by Armstrong-Carter and Telzer (2021), positive eco-friendly actions and climate activism should be considered important forms of prosocial behavior in the 21st century, because these behaviors contribute positively to the planet and the lives of others.

6. Conclusions and future directions

This review set out to examine the question: When and how do changes in emotional reactivity and socio-cognitive development result in opportunities for prosocial development and which factors facilitate opportunities for positive, prosocial development? We demonstrated that prosocial behavior is a multidimensional construct (Carlo & Padilla-Walker, 2020), where age patterns are related to the emergence of developmental processes as well as increasing sensitivity to social and contextual factors.

In terms of developmental processes, we reviewed evidence showing socio-affective neural reactivity to rewards for close others, consistent with the hypothesis that valuing rewards for self and others relies on the same neural system, including the dopamine-rich ventral striatum and the vmPFC (Telzer, 2016). There is some evidence for heightened activity in the ventral striatum in mid-adolescence for vicarious rewards for close friends and for mothers (Braams & Crone, 2017; Schreuders et al., 2021), but these developmental patterns are not consistent across studies (Brandner et al., 2021) and warrant further investigation, including the relation with warmth of the relationship or influences of childhood experiences (Decety & Holvoet, 2021). Socio-cognitive developmental processes such as perspective taking and goal flexibility show consistently protracted changes during adolescence; including changes in the medial PFC, TPJ and dlPFC (Do et al., 2019; van de Groep et al., 2022). Together, the reviewed studies provide consistent support for a model showing dynamic changes in socio-affective and socio-cognitive changes during adolescence.

We argue that the emergence of these developmental processes interacts with contextual processes. First, we showed that individuals give more to in-group compared to out-group members, and more to close, familiar others compared to unknown others (van de Groep, Zanolie, & Crone, 2020b). Although in-group biases materialize early in life, recent studies showed that differentiation between targets further increases over the course of adolescence (Guroglu et al., 2014). Second, from childhood to adolescence, individuals move from simple decisions such as equal splits in childhood, to more nuanced decisions that incorporate the social situation and require perspective taking in adolescence, such as differentiating between the target of giving (Do & Telzer, 2019). Third, giving decisions are influenced by whether individuals are observed by others (audience effects) (van de Groep, Zanolie, & Crone, 2020a) and these influences may make adolescents also sensitive to prosocial peer interventions (Van Hoorn, Van Dijk, Guroglu, & Crone, 2016; Veenstra & Laninga-Wijnen, 2021). Future studies should examine how this working model may aid in developing interventions that shape prosocial behavior toward family, school or neighborhood. These interventions should also examine how developing prosocial behavior may influence the need of adolescents to contribute (Fuligni, 2019) and the need to experience purpose and meaning (Yeager et al., 2018).

There are several methodological questions that require more emphasis in future research. The reviewed studies examined adolescents in relatively isolated lab environments. In future studies it will be important to test actual online or offline interactions. Further, we used a decomposition method, but in future studies it will be interesting to test the multidimensionality of prosocial behavior to examine how (latent) processes interact with each other within individuals. We showed in Box 1 an example of a large longitudinal study from our group that allows for the test of these relations. Finally, it should be noted that we reviewed relatively simple economic games formats where participants were confronted with resources to share with others. In future research, it should be examined how the various contexts test not only the motivation ("willingness") and developmental processes ("capacity"), but also the resources to give to others ("possibility") of adolescents.

Future studies should examine prosocial behavior in relation to the challenges that adolescents face today and with the tools that adolescent use to influence their environment. Adolescents face complex societal challenges that require a vision that goes beyond country borders (climate changes, pandemics, global inequalities) (Orben, Tomova, & Blakemore, 2020) but adolescents also have the tools to influence at larger scales using digital networks (Armstrong-Carter & Telzer, 2021). It was previously found that reward sensitivity in mid-adolescence predicts both intentions to be rebellious as well as prosocial behavior (Blankenstein et al., 2020). Possibly, adolescents, who traditionally have been categorized as a risk group, may have a large drive to influence society by combining rebelliousness with prosocial behaviors, providing a possible pathway to a socially cohesive and cooperative society with the ability to tackle tomorrow's global challenges (Do, Guassi Moreira, & Telzer, 2017).

Acknowledgments

E.A.C. is supported by an "innovative ideas" grant of the European Research Council (ERC CoG PROSOCIAL 681632 to E.A.C.). We thank Iris Langereis for helping with the illustrations.

Declaration of interest

None.

Appendix 1

		Longitudinal neuroimaging study						Intervention study			
		T1 - Adolescents	T1 - Parent study	T1.5	T2	COVID-19 daily diary study	T3	T1	T2	T3	
		May - Oct 2018	Oct 2018 - March 2019	Every 2 months between T1 and T2	Aug 2019 - Jan 2020	March - April 2020	October 2021 -	October 2021	ongoing	ongoing	
Measurements	**Respondents**										**Comments**
Prosocial behavior											
Neurobiological measures											
Cooperation	Child, Parent(s)	x	x								
Giving	Child	x			x		x				
Trust and Reciprocity	Child						x				
Behavioral experiments											
Social Dilemma	Child, Parent(s)	x	x				x				
Naturalistic Helping Task	Child	x									
Charity Dictator Game	Child	x			x		x				
Societal Trust Game	Child									x	
Pandemic Dictator Game	Child					x				x	Five targets: unfamiliar peer, friend, individual with poor immune system, individual with COVID-19, medical doctor
Questionnaires											
Opportunities for Prosocial Actions (OPA)	Child	x		5x	x	15x	x	x	x	x	At COVID-19 daily diary study, only the subscale emotional support was administered. At the intervention study, a revised version was administered
Prosocial Tendencies Measure Revised (PTM-R)	Child, Parent(s)	x	x		x	15x	x	x			At COVID-19 daily diary study, only the subscales Dire and Altruism were administered. At the intervention study, only the subscale Altruism was administered.
Social Value Orientation (SVO)	Child, Parent(s)	x	x		x	2x	x				
Digital Prosocial Behavior	Child				x		x				
Contributions to society - General	Child				x	x	x	x	x	x	At the intervention study, a revised version was administered
Contributions to society - Crisis	Child					x					
Pandemic Questionnaire - Helping	Child					x					
Online Prosocial Behavior Scale	Child						x				
Need for Useful Contributions - Revised measure	Child						x	x	x	x	
Motives for Prosocial Behavior	Child							x	x	x	
Personal Characteristics											
Neurobiological measures											
Delay Discounting For Self and Other	Child				x						
Behavioral experiments											
Stop-Signal Task (Impulsivity)	Child	x			x		x				
Balloon Analogue Risk Task (Risk-taking)	Child	x									
WISC/WAIS similarities (verbal IQ)	Child	x									
WISC/WAIS block design (Visual/Spatial IQ)	Child	x									
Three Minute Test (Reading Fluency)	Child	x									
Questionnaires											
Demographics	Child, Parent(s)	x	x		x	x	x				
Self Control Scale	Child	x			x		x				
Adolescent Risk-taking Questionnaire	Child	x			x		x				
Positive Risk-taking Scale	Child				x		x				
Rosenberg Self-esteem Scale	Child	x			x		x				
Interpersonal Reactivity Index	Child	x			x	4x	x	x	x	x	Only subscales Empathic Concern and Perspective Taking were administered. At the intervention study, the subscale empathic concern was only administered at T1
Social Desirability Response Set	Child	x			x	x	x				
Emotion Awareness Questionnaire (revised)	Child	x			x		x				
HEXACO	Child, Parent(s)	x	x								
Religious Background and Behaviors Questionnaire (RBBQ)	Child, Parent(s)	x	x		x		x				
Adult Temperament Questionnaire (ATQ)	Parent	x	x		x		x				
Pandemic Questionnaire - Perceived Risk	Child					15x					
Pandemic Questionnaire - Selfish behavior	Child					15x					
Risk Taking (1 question, scale 0 - 100)	Child					2x					
Childhood Narcissism Scale	Child						x				
Feeling useful	Child				x		x				
The Brief Sensation-Seeking Scale	Child							x			
Social Tendencies											
Questionnaires											
Social Reward Questionnaire (adolescent version)	Child	x			x		x	x			At the intervention study, only the subscale Prosocial Interactions was administered
Social Comparison Orientation	Child	x			x						
Intrasexual Competition	Child	x									
Olweus Bully/Victimization Questionnaire	Child	x									
Inclusion of Other in Self	Child	x	x		x		x				
Experiences in Close Relationships Revised	Child	x	x	5x	x	x	x				
Friendship Quality Scale	Child	x			x		x				
Romantic Relationships, Gender Identity, and Sexual Orientation	Child	x			x						
Self in a Social Context-Social Connectedness Scale	Child							x	x	x	
Strenghts and Difficulties Questionnaire	Parent(s)	x			x						
Social Responsiveness Scale	Parent(s)	x			x						
Family-related Measures											
Sibling Relationship Questionnaire	Child	x									
Family Adaptability and Cohesion Evaluation Scale (FACES)	Child, Parent(s)	x	x		x		x				
Leuven Adolescent Perceived Parenting Scale (LAPPS)	Child, Parent(s)	x	x		x		x				
Physical/physiological measures											
Neurobiological measures											
Structural MRI scan	Child, Parent(s)	x			x		x				
Resting State MRI scan	Child, Parent(s)	x			x		x				
Diffusion Tensor Imaging scan	Child, Parent(s)	x			x						
Hormonal measures											
Testosterone	Child, Parent(s)	x			x		x				
Estradiol	Child, Parent(s)	x			x		x				
Oxytocin	Child, Parent(s)	x									
DHEA	Child, Parent(s)	x			x		x				
Cortisol	Child, Parent(s)	x									
Progesterone	Child, Parent(s)	x									
Questionnaires											
Pubertal Development Scale	Child	x			x		x				
Menstrual Cycle Questionnaire	Child	x			x						
Life Events											
Questionnaires											
LEC-5 (Life Event Checklist)	Parent(s)		x								
Child and Adolescent Survey of Experiences (CASE)	Child			5x	x						
Pandemic Questionnaire - Home Isolation	Child					15x					
Pandemic Experiences	Child						x				
Mood/Wellbeing											
Questionnaires											
Profile of Mood Scale (POMS)	Child			5x		15x		x	x	x	At COVID-19 daily diary study, only the subscales Vigor and Tension were administered. At the intervention study, only the subscales Vigor and Depression were administered
Pandemic Questionnaire - Stress	Child					15x					
Multidimensional Wellbeing Paradigm	Child						x	x	x	x	At the intervention study, a shortened version was administered

References

Achterberg, M., van Duijvenvoorde, A. C., Bakermans-Kranenburg, M. J., & Crone, E. A. (2016). Control your anger! The neural basis of aggression regulation in response to negative social feedback. *Social Cognitive and Affective Neuroscience*, *11*(5), 712–720. https://doi.org/10.1093/scan/nsv154.

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews. Neuroscience*, 7(4), 268–277. https://doi.org/10.1038/nrn1884.

Andolina, M. W., Jenkins, K., Zukin, C., & Keeter, S. (2003). Habits from home, lessons from school: Influences on youth civic engagement. *Political Science and Politics*, *36*(3), 275–280. https://doi.org/10.1017/S104909650300221X.

Armstrong-Carter, E., & Telzer, E. H. (2021). Advancing measurement and research on Youths' prosocial behavior in the digital age. *Child Development Perspectives*, *15*(1), 31–36. https://doi.org/10.1111/cdep.1239.

Arnett, J. J., Zukauskiene, R., & Sugimura, K. (2014). The new life stage of emerging adulthood at ages 18-29 years: Implications for mental health. *Lancet Psychiatry*, *1*(7), 569–576. https://doi.org/10.1016/S2215-0366(14)00080-7.

Ballard, P. J., Cohen, A. K., & Littenberg-Tobias, J. (2016). Action civics for promoting civic development: Main effects of program participation and differences by project characteristics. *American Journal of Community Psychology*, *58*, 377–390. https://doi.org/10.1002/ajcp.12103.

Barry, C. M., & Wentzel, K. R. (2006). Friend influence on prosocial behavior: The role of motivational factors and friendship characteristics. *Developmental Psychology*, *42*(1), 153–163. https://doi.org/10.1037/0012-1649.42.1.153.

Baumsteiger, R. (2019). What the world needs now: An intervention for promoting prosocial behavior. *Basic and Applied Social Psychology*, *41*(4), 215–229. https://doi.org/10.1080/01973533.2019.1639507.

Becht, A. I., Wierenga, L. M., Mills, K. L., Meuwese, R., van Duijvenvoorde, A., Blakemore, S. J., et al. (2020). Beyond the average brain: Individual differences in social brain development are associated with friendship quality. *Social Cognitive and Affective Neuroscience*, *16*(3), 292–301. https://doi.org/10.1093/scan/nsaa166.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, *10*, 122–142.

Blakemore, S. J., & Mills, K. L. (2014). Is adolescence a sensitive period for sociocultural processing? *Annual Review of Psychology*, *65*, 187–207. https://doi.org/10.1146/annurev-psych-010213-115202.

Blankenstein, N. E., Telzer, E. H., Do, K. T., van Duijvenvoorde, A. C. K., & Crone, E. A. (2020). Behavioral and neural pathways supporting the development of prosocial and risk-taking behavior across adolescence. *Child Development*, *91*(3), e665–e681. https://doi.org/10.1111/cdev.13292.

Blote, A. W., Kint, M. J., Miers, A. C., & Westenberg, P. M. (2009). The relation between public speaking anxiety and social anxiety: A review. *Journal of Anxiety Disorders*, *23*(3), 305–313. https://doi.org/10.1016/j.janxdis.2008.11.007.

Boyes, M. E., & French, D. J. (2009). Having a cyberball: Using a ball-throwing game as an experimental social stressor to examine the relationship between neuroticism and coping. *Personality and Individual Differences*, *47*(5), 396–401.

Braams, B. R., & Crone, E. A. (2017). Peers and parents: A comparison between neural activation when winning for friends and mothers in adolescence. *Social Cognitive and Affective Neuroscience*, *12*(3), 417–426. https://doi.org/10.1093/scan/nsw136.

Brandner, P., Guroglu, B., & Crone, E. A. (2020). I am happy for us: Neural processing of vicarious joy when winning for parents versus strangers. *Cognitive, Affective, & Behavioral Neuroscience*, *20*(6), 1309–1322. https://doi.org/10.3758/s13415-020-00839-9.

Brandner, P., Guroglu, B., van de Groep, S., Spaans, J. P., & Crone, E. A. (2021). Happy for us not them: Differences in neural activation in a vicarious reward task between family and strangers during adolescent development. *Developmental Cognitive Neuroscience*, *51*, 100985. https://doi.org/10.1016/j.dcn.2021.100985.

Brownell, C. A., Svetlova, M., Anderson, R., Nichols, S. R., & Drummond, J. (2013). Socialization of early prosocial behavior: Parents' talk about emotions is associated with sharing and helping in toddlers. *Infancy*, *18*, 91–119. https://doi.org/10.1111/j.1532-7078.2012.00125.x.

Burke, S. M., van de Groep, S., Brandner, P., & Crone, E. A. (2020). Neurocognitive developmental changes in trust and reciprocity across adolescence. In K. Cohen Kadosh (Ed.), *Developmental cognitive neuroscience*. Oxford: Oxford University Press.

Caprara, G. V., Kanacri, B. P. L., Gerbino, M., Zuffianò, A., Alessandri, G., Vecchio, G., et al. (2014). Positive effects of promoting prosocial behavior in early adolescence: Evidence from a school-based intervention. *International Journal of Behavioral Development*, *38*(4), 386–396. https://doi.org/10.1177/0165025414531464.

Caprara, G. V., Luengo Kanacri, B. P., Zuffianò, A., Gerbino, M., & Pastorelli, C. (2015). Why and how to promote adolescents' prosocial behaviors: Direct, mediated and moderated effects of the CEPIDEA school-based program. *Journal of Youth and Adolescence*, *44*(12), 2211–2229. https://doi.org/10.1007/s10964-015-0293-1.

Carlo, G., & Padilla-Walker, L. (2020). Adolescents' prosocial behaviour though a multidimensional and multicultural lens. *Child Development Perspectives*, *14*(4), 265–272.

Carter, R. M., & Huettel, S. A. (2013). A nexus model of the temporal-parietal junction. *Trends in Cognitive Sciences*, *17*(7), 328–336. https://doi.org/10.1016/j.tics.2013.05.007.

Casey, B. J. (2015). Beyond simple models of self-control to circuit-based accounts of adolescent behavior. *Annual Review of Psychology*, *66*, 295–319. https://doi.org/10.1146/annurev-psych-010814-015156.

Chein, J., Albert, D., O'Brien, L., Uckert, K., & Steinberg, L. (2011). Peers increase adolescent risk taking by enhancing activity in the brain's reward circuitry. *Developmental Science*, *14*(2), F1–10. https://doi.org/10.1111/j.1467-7687.2010.01035.x.

Choukas-Bradley, S., Giletta, M., Cohen, G. L., & Prinstein, M. J. (2015). Peer influence, peer status, and prosocial behavior: An experimental investigation of peer socialization of adolescents' intentions to volunteer. *Journal of Youth and Adolescence*, *44*(12), 2197–2210. https://doi.org/10.1007/s10964-015-0373-2.

Cohen, J. R., Asarnow, R. F., Sabb, F. W., Bilder, R. M., Bookheimer, S. Y., Knowlton, B. J., et al. (2010). A unique adolescent response to reward prediction errors. *Nature Neuroscience*, *13*(6), 669–671. https://doi.org/10.1038/nn.2558.

Crone, E. A., & Dahl, R. E. (2012). Understanding adolescence as a period of social-affective engagement and goal flexibility. *Nature Reviews. Neuroscience*, *13*(9), 636–650. https://doi.org/10.1038/nrn3313.

Crone, E. A., & Fuligni, A. J. (2020). Self and others in adolescence. *Annual Review of Psychology*, *71*, 447–469. https://doi.org/10.1146/annurev-psych-010419-050937.

Curry, O. S., Rowlans, L. A., Van Lissa, C. J., Zlotowitz, S., McAlaney, J., & Whitehouse, H. (2018). Happy to help? A systematic review and meta-analysis of the effects of performing acts of kindness on the well-being of the actor. *Journal of Experimental Social Psychology*, *76*, 320–329. https://doi.org/10.1016/j.jesp.2018.02.014.

Cutler, J., & Campbell-Meiklejohn, D. (2019). A comparative fMRI meta-analysis of altruistic and strategic decisions to give. *NeuroImage*, *184*, 227–241. https://doi.org/10.1016/j.neuroimage.2018.09.009.

Dahl, R. E., Allen, N. B., Wilbrecht, L., & Suleiman, A. B. (2018). Importance of investing in adolescence from a developmental science perspective. *Nature*, *554*(7693), 441–450. https://doi.org/10.1038/nature25770.

Decety, J., & Holvoet, C. (2021). The emergence of empathy: A developmental neuroscience perspective. *Developmental Review*, *62*, 100999.

Delgado, M. R., Beer, J. S., Fellows, L. K., Huettel, S. A., Platt, M. L., Quirk, G. J., et al. (2016). Viewpoints: Dialogues on the functional role of the ventromedial prefrontal cortex. *Nature Neuroscience*, *19*(12), 1545–1552. https://doi.org/10.1038/nn.4438.

Do, K. T., Guassi Moreira, J. F., & Telzer, E. H. (2017). But is helping you worth the risk? Defining prosocial risk taking in adolescence. *Developmental Cognitive Neuroscience*, *25*, 260–271. https://doi.org/10.1016/j.dcn.2016.11.008.

Do, K. T., McCormick, E. M., & Telzer, E. H. (2019). The neural development of prosocial behavior from childhood to adolescence. *Social Cognitive and Affective Neuroscience*, *14*(2), 129–139. https://doi.org/10.1093/scan/nsy117.

Do, K. T., & Telzer, E. H. (2019). Corticostriatal connectivity is associated with the reduction of intergroup bias and greater impartial giving in youth. *Developmental Cognitive Neuroscience*, *37*, 100628. https://doi.org/10.1016/j.dcn.2019.100628.

Dumontheil, I., Apperly, I. A., & Blakemore, S. J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science*, *13*(2), 331–338. https://doi.org/10.1111/j.1467-7687.2009.00888.x.

Ellis, B. J., Volk, A. A., Gonzalez, J. M., & Embry, D. D. (2016). The meaningful roles intervention: An evolutionary approach to reducing bullying and increasing prosocial behavior. *Journal of Research on Adolescence*, *26*(4), 622–637. https://doi.org/10.1111/jora.12243.

Engel, C. (2011). Dictator games: A meta study. *Experimental Economics*, *14*(4), 583–610.

Ernst, M. (2014). The triadic model perspective for the study of adolescent motivated behavior. *Brain and Cognition*, *89*, 104–111. https://doi.org/10.1016/j.bandc.2014.01.006.

Etxebarria, I., Apodaka, P., Eceiza, A., Ortiz, M. J., Fuentes, M. J., & Lopez, F. (1994). Design and evaluation of a programme to promote prosocial-altruistic behaviour in the school. *Journal of Moral Education*, *23*(4), 409–425. https://doi.org/10.1080/0305724940230403.

Feng, C., Luo, Y. J., & Krueger, F. (2015). Neural signatures of fairness-related normative decision making in the ultimatum game: A coordinate-based meta-analysis. *Human Brain Mapping*, *36*(2), 591–602. https://doi.org/10.1002/hbm.22649.

Fett, A. K., Gromann, P. M., Giampietro, V., Shergill, S. S., & Krabbendam, L. (2014). Default distrust? An fMRI investigation of the neural development of trust and cooperation. *Social Cognitive and Affective Neuroscience*, *9*(4), 395–402. https://doi.org/10.1093/scan/nss144.

Fett, A. K., Shergill, S. S., Gromann, P. M., Dumontheil, I., Blakemore, S. J., Yakub, F., et al. (2014). Trust and social reciprocity in adolescence—A matter of perspective-taking. *Journal of Adolescence*, *37*(2), 175–184. https://doi.org/10.1016/j.adolescence.2013.11.011.

Foulkes, L., Leung, J. T., Fuhrmann, D., Knoll, L. J., & Blakemore, S. J. (2018). Age differences in the prosocial influence effect. *Developmental Science*, *21*(6), e12666. https://doi.org/10.1111/desc.12666.

Fu, X., Padilla-Walker, L. M., & Brown, M. N. (2017). Longitudinal relations between adolescents' self-esteem and prosocial behavior toward strangers, friends and family. *Journal of Adolescence*, *57*, 90–98. https://doi.org/10.1016/j.adolescence.2017.04.002.

Fuligni, A. J. (2019). The need to contribute in adolescence. *Perspectives on Psychological Science*, *14*(3), 331–343. https://doi.org/10.1177/1745691618805437.

Fuligni, A. J. (2020). Is there inequality in what adolescents can give as well as receive? *Current Directions in Psychological Science*, *29*(4), 405–411. https://doi.org/10.1177/0963721420917738.

Gagné, M. (2003). The role of autonomy support and autonomy orientation in prosocial behavior engagement. *Motivation and Emotion*, *27*(3), 199–223.

Galvan, A. (2010). Adolescent development of the reward system. *Frontiers in Human Neuroscience*, *4*, 6. https://doi.org/10.3389/neuro.09.006.2010.

Galvan, A., Hare, T. A., Parra, C. E., Penn, J., Voss, H., Glover, G., et al. (2006). Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents. *The Journal of Neuroscience*, *26*(25), 6885–6892. https://doi.org/10.1523/JNEUROSCI.1062-06.2006.

Garaigordobil, M. (2008). Assessment of the effects of a cooperative play programme for children aged 10–11 years on social adaptation and on the perception that parents, teachers and peers have of children's prosocial behaviours. *Infancia y Aprendizaje*, *31*(3), 303–318. https://doi.org/10.1174/021037008785702974.

Gilmore, J. H., Knickmeyer, R. C., & Gao, W. (2018). Imaging structural and functional brain development in early childhood. *Nature Reviews. Neuroscience*, *19*(3), 123–137. https://doi.org/10.1038/nrn.2018.1.

Goddings, A. L., Mills, K. L., Clasen, L. S., Giedd, J. N., Viner, R. M., & Blakemore, S. J. (2014). The influence of puberty on subcortical brain development. *NeuroImage*, *88*, 242–251. https://doi.org/10.1016/j.neuroimage.2013.09.073.

Gummerum, M., Hanoch, Y., & Keller, M. (2008). When child development meets economic game theory: An interdisciplinary approach to investigating social development. *Human Development*, *51*(4), 235–261.

Guroglu, B., van den Bos, W., & Crone, E. A. (2014). Sharing and giving across adolescence: An experimental study examining the development of prosocial behavior. *Frontiers in Psychology*, *5*, 291. https://doi.org/10.3389/fpsyg.2014.00291.

Guroglu, B., van den Bos, W., van Dijk, E., Rombouts, S. A., & Crone, E. A. (2011). Dissociable brain networks involved in development of fairness considerations: Understanding intentionality behind unfairness. *NeuroImage*, *57*(2), 634–641. https://doi.org/10.1016/j.neuroimage.2011.04.032.

Haber, S. N., & Knutson, B. (2010). The reward circuit: Linking primate anatomy and human imaging. *Neuropsychopharmacology*, *35*(1), 4–26. https://doi.org/10.1038/npp.2009.129.

Hackel, L. M., Zaki, J., & Van Bavel, J. J. (2017). Social identity shapes social valuation: Evidence from prosocial behavior and vicarious reward. *Social Cognitive and Affective Neuroscience*, *12*(8), 1219–1228. https://doi.org/10.1093/scan/nsx045.

Harbaugh, W. T., Mayr, U., & Burghart, D. R. (2007). Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science*, *316*(5831), 1622–1625. https://doi.org/10.1126/science.1140738.

Herting, M. M., Johnson, C., Mills, K. L., Vijayakumar, N., Dennison, M., Liu, C., et al. (2018). Development of subcortical volumes across adolescence in males and females: A multisample study of longitudinal changes. *NeuroImage*, *172*, 194–205. https://doi.org/10.1016/j.neuroimage.2018.01.020.

Lahno, B. (1995). Trust, reputation, and exit in exchange relationships. *Journal of Conflict Resolution*, *39*(3), 495–510.

Larson, R. W., Moneta, G., Richards, M. H., & Wilson, S. (2002). Continuity, stability, and change in daily emotional experience across adolescence. *Child Development*, *73*(4), 1151–1165.

Lee, F. S., Heimer, H., Giedd, J. N., Lein, E. S., Sestan, N., Weinberger, D. R., et al. (2014). Mental health. Adolescent mental health—Opportunity and obligation. *Science*, *346*(6209), 547–549. https://doi.org/10.1126/science.1260497.

Lee, S. J., Zhang, J., Neale, M. C., Styner, M., Zhu, H., & Gilmore, J. H. (2019). Quantitative tract-based white matter heritability in 1- and 2-year-old twins. *Human Brain Mapping*, *40*(4), 1164–1173. https://doi.org/10.1002/hbm.24436.

Lieberman, M. D., Straccia, M. A., Meyer, M. L., Du, M., & Tan, K. M. (2019). Social, self, (situational), and affective processes in medial prefrontal cortex (MPFC): Causal, multivariate, and reverse inference evidence. *Neuroscience and Biobehavioral Reviews*, *99*, 311–328. https://doi.org/10.1016/j.neubiorev.2018.12.021.

Lockwood, P. L., Apps, M. A., Valton, V., Viding, E., & Roiser, J. P. (2016). Neurocomputational mechanisms of prosocial learning and links to empathy. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(35), 9763–9768. https://doi.org/10.1073/pnas.1603198113.

Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*, *453*(7197), 869–878.
Luo, J. (2018). The neural basis of and a common neural circuitry in different types of pro-social behavior. *Frontiers in Psychology*, *9*, 859. https://doi.org/10.3389/fpsyg.2018.00859.
Malonda, E., Llorca, A., Mesurado, B., Samper, P., & Mestre, M. V. (2019). Parents or peers? Predictors of prosocial behavior and aggression: A longitudinal study. *Frontiers in Psychology*, *10*, 2379. https://doi.org/10.3389/fpsyg.2019.02379.
Malti, T., Averdijk, M., Zuffiano, A., Ribeaud, D., Betts, L. R., Rotenberg, K. J., et al. (2015). Children's trust and the development of prosocial behavior. *International Journal of Behavioral Development*, *40*(3), 262–270.
McGinley, M., Lipperman-Kreda, S., Byrnes, H. F., & Carlo, G. (2010). Parental, social and dispositional pathways to Israeli adolescents' volunteering. *Journal of Applied Developmental Psychology*, *31*(5), 386–394. https://doi.org/10.1016/j.appdev.2010.06.001.
Mesurado, B., Distefano, M. J., Robiolo, G., & Richaud, M. C. (2019). The Hero program: Development and initial validation of an intervention program to promote prosocial behavior in adolescents. *Journal of Social and Personal Relationships*, *36*(8), 2566–2584. https://doi.org/10.1177/0265407518793224.
Mesurado, B., Guerra, P., Richaud, M. C., & Rodrigues, L. M. (2019). Effectiveness of prosocial behavior interventions: A meta-analysis. In *Psychiatry and neuroscience update* (pp. 259–271). Springer.
Mesurado, B., & Richaud, M. C. (2018). The relationship between parental variables. *Journal of Happiness Studies*, *18*(3), 843–860. https://doi.org/10.1007/s10902-016-9748-7.
Meuwese, R., Crone, E. A., de Rooij, M., & Guroglu, B. (2015). Development of equity preferences in boys and girls across adolescence. *Child Development*, *86*(1), 145–158. https://doi.org/10.1111/cdev.12290.
Mills, K. L., Goddings, A.-L., Clasen, L. S., Giedd, J. N., & Blakemore, S.-J. (2014). The developmental mismatch in structural brain maturation during adolescence. *Developmental Neuroscience*, *36*, 147–160.
Mills, K. L., Goddings, A. L., Herting, M. M., Meuwese, R., Blakemore, S. J., Crone, E. A., et al. (2016). Structural brain development between childhood and adulthood: Convergence across four longitudinal samples. *NeuroImage*, *141*, 273–281. https://doi.org/10.1016/j.neuroimage.2016.07.044.
Morelli, S. A., Knutson, B., & Zaki, J. (2018). Neural sensitivity to personal and vicarious reward differentially relates to prosociality and well-being. *Social Cognitive and Affective Neuroscience*, *13*(8), 831–839. https://doi.org/10.1093/scan/nsy056.
Orben, A., Tomova, L., & Blakemore, S. J. (2020). The effects of social deprivation on adolescent development and mental health. *The Lancet Child & Adolescent Health*, *4*(8), 634–640. https://doi.org/10.1016/S2352-4642(20)30186-3.
Padilla-Walker, L., Carlo, G., & Memmott-Elison, M. K. (2018). Longitudinal change in adolescents' prosocial behavior toward strangers, friends, and family. *Journal of Research on Adolescence*, *28*(3), 698–710. https://doi.org/10.1111/jora.12362.
Paus, T. (2010). Growth of white matter in the adolescent brain: Myelin or axon? *Brain and Cognition*, *72*(1), 26–35. https://doi.org/10.1016/j.bandc.2009.06.002.
Paus, T., Keshavan, M., & Giedd, J. N. (2008). Why do many psychiatric disorders emerge during adolescence? *Nature Reviews. Neuroscience*, *9*(12), 947–957. https://doi.org/10.1038/nrn2513.
Pope, A., Stolte, L., & Cohen, A. (2011). Closing the civic engagement gap: The potential of action civics. *Social Education*, *75*(5), 265–268.
Riem, M. M., Bakermans-Kranenburg, M. J., Huffmeijer, R., & van IJzendoorn, M. H. (2013). Does intranasal oxytocin promote prosocial behavior to an excluded fellow

player? A randomized-controlled trial with Cyberball. *Psychoneuroendocrinology*, *38*(8), 1418–1425. https://doi.org/10.1016/j.psyneuen.2012.12.023.

Rilling, J., Gutman, D., Zeh, T., Pagnoni, G., Berns, G., & Kilts, C. (2002). A neural basis for social cooperation. *Neuron*, *35*(2), 395–405. https://doi.org/10.1016/s0896-6273(02)00755-9.

Rilling, J. K., & Sanfey, A. G. (2011). The neuroscience of social decision-making. *Annual Review of Psychology*, *62*, 23–48. https://doi.org/10.1146/annurev.psych.121208.131647.

Sabato, H., Abraham, Y., & Kogut, T. (2021). Too lonely to help: Early adolescents' social connections and willingness to help during COVID-19 lockdown. *Journal of Research on Adolescence*, *31*(3), 764–779. https://doi.org/10.1111/jora.12655.

Samuels, W. E. (2018). Nurturing kindness naturally: A humane education program's effect on the prosocial behavior of first and second graders across China. *International Journal of Educational Research*, *91*, 49–64. https://doi.org/10.1016/j.ijer.2018.08.001.

Schonert-Reichl, K. A., Smith, V., Zaidman-Zait, A., & Hertzman, C. (2012). Promoting children's prosocial behaviors in school: Impact of the "roots of empathy" program on the social and emotional competence of school-aged children. *School Mental Health*, *4*(1), 1–21. https://doi.org/10.1007/s12310-011-9064-7.

Schreuders, E., Braams, B. R., Crone, E. A., & Guroglu, B. (2021). Friendship stability in adolescence is associated with ventral striatum responses to vicarious rewards. *Nature Communications*, *12*(1), 313. https://doi.org/10.1038/s41467-020-20042-1.

Schreuders, E., Klapwijk, E. T., Will, G. J., & Guroglu, B. (2018). Friend versus foe: Neural correlates of prosocial decisions for liked and disliked peers. *Cognitive, Affective, & Behavioral Neuroscience*, *18*(1), 127–142. https://doi.org/10.3758/s13415-017-0557-1.

Shin, J., & Lee, B. (2021). The effects of adolescent prosocial behavior interventions: A meta-analytic review. *Asia Pacific Education Review*, *22*(3), 565–577. https://doi.org/10.1007/s12564-021-09691-z.

Shirtcliff, E. A., Dahl, R. E., & Pollak, S. D. (2009). Pubertal development: Correspondence between hormonal and physical development. *Child Development*, *80*(2), 327–337. https://doi.org/10.1111/j.1467-8624.2009.01263.x.

Silverman, M. H., Jedd, K., & Luciana, M. (2015). Neural networks involved in adolescent reward processing: An activation likelihood estimation meta-analysis of functional neuroimaging studies. *NeuroImage*, *122*, 427–439. https://doi.org/10.1016/j.neuroimage.2015.07.083.

Smetana, J. G., & Metzger, A. (2005). Family and religious antecedents of civic involvement in middle class African American late adolescents. *Journal of Research on Adolescence*, *15*(3), 325–352. https://doi.org/10.1111/j.1532-7795.2005.00099.x.

Spaans, J. P., Peters, S., & Crone, E. A. (2019). Neural reward-related reactions to monetary gains for self and charity. *Cognitive, Affective, & Behavioral Neuroscience*, *19*(4), 845–858. https://doi.org/10.3758/s13415-018-00672-1.

Spaans, J. P., Peters, S., & Crone, E. A. (2020). Neural reward related-reactions to monetary gains for self and charity are associated with donating behavior in adolescence. *Social Cognitive and Affective Neuroscience*, *15*(2), 151–163. https://doi.org/10.1093/scan/nsaa027.

Spaans, J. P., Will, G. J., van Hoorn, J., & Guroglu, B. (2019). Turning a blind eye? Punishment of friends and unfamiliar peers after observed exclusion in adolescence. *Journal of Research on Adolescence*, *29*(2), 508–522. https://doi.org/10.1111/jora.12401.

Steinbeis, N., Bernhardt, B. C., & Singer, T. (2012). Impulse control and underlying functions of the left DLPFC mediate age-related and age-independent individual differences in strategic social behavior. *Neuron*, *73*(5), 1040–1051. https://doi.org/10.1016/j.neuron.2011.12.027.

Steinberg, L. (2008). *Adolescence*. McGraw-Hill.

Stern, J. A., & Cassidy, J. (2018). Empathy from infancy to adolescence: An attachment perspective on the development of individual differences. *Developmental Review*, *47*, 1–22.

Sweijen, S. W., van de Groep, S., Green, K. H., te Brinke, L., Buijzen, M., de Leeuw, R. N. H., & Crone, E. A., n.d., Daily prosocial actions during the COVID-19 pandemic contribute to giving behavior in adolescence. https://psyarxiv.com/7cptg/, (preprint).

Tabibnia, G., Satpute, A. B., & Lieberman, M. D. (2008). The sunny side of fairness: Preference for fairness activates reward circuitry (and disregarding unfairness activates self-control circuitry). *Psychological Science*, *19*(4), 339–347. https://doi.org/10.1111/j.1467-9280.2008.02091.x.

Tamir, D. I., & Hughes, B. L. (2018). Social rewards: From basic social building blocks to complex social behavior. *Perspectives on Psychological Science*, *13*(6), 700–717. https://doi.org/10.1177/1745691618776263.

Tamnes, C. K., Herting, M. M., Goddings, A. L., Meuwese, R., Blakemore, S. J., Dahl, R. E., et al. (2017). Development of the cerebral cortex across adolescence: A multisample study of inter-related longitudinal changes in cortical volume, surface area, and thickness. *The Journal of Neuroscience*, *37*(12), 3402–3412. https://doi.org/10.1523/JNEUROSCI.3302-16.2017.

Tashjian, S. M., Weissman, D. G., Guyer, A. E., & Galvan, A. (2018). Neural response to prosocial scenes relates to subsequent giving behavior in adolescents: A pilot study. *Cognitive, Affective, & Behavioral Neuroscience*, *18*(2), 342–352. https://doi.org/10.3758/s13415-018-0573-9.

te Brinke, L. W., van de Groep, S., van der Cruijsen, R., & Crone, E. A., n.d. Variability and Change in Adolescents' Prosocial Behavior Across Multiple Time Scales. doi: 10.31234/osf.io/cqntp (web archive link), (preprint).

Telzer, E. H. (2016). Dopaminergic reward sensitivity can promote adolescent health: A new perspective on the mechanism of ventral striatum activation. *Developmental Cognitive Neuroscience*, *17*, 57–67.

Telzer, E. H., Fuligni, A. J., Lieberman, M. D., & Galvan, A. (2013). Ventral striatum activation to prosocial rewards predicts longitudinal declines in adolescent risk taking. *Developmental Cognitive Neuroscience*, *3*, 45–52. https://doi.org/10.1016/j.dcn.2012.08.004.

Telzer, E. H., Fuligni, A. J., Lieberman, M. D., & Galvan, A. (2014). Neural sensitivity to eudaimonic and hedonic rewards differentially predict adolescent depressive symptoms over time. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(18), 6600–6605. https://doi.org/10.1073/pnas.1323014111.

Telzer, E. H., Ichien, N., & Qu, Y. (2015). The ties that bind: Group membership shapes the neural correlates of in-group favoritism. *NeuroImage*, *115*, 42–51. https://doi.org/10.1016/j.neuroimage.2015.04.035.

Telzer, E. H., Masten, C. L., Berkman, E. T., Lieberman, M. D., & Fuligni, A. J. (2010). Gaining while giving: An fMRI study of the rewards of family assistance among white and Latino youth. *Social Neuroscience*, *5*(5–6), 508–518. https://doi.org/10.1080/17470911003687913.

Tousignant, B., Eugene, F., Sirois, K., & Jackson, P. L. (2018). Difference in neural response to social exclusion observation and subsequent altruism between adolescents and adults. *Neuropsychologia*, *116*(Pt. A), 15–25. https://doi.org/10.1016/j.neuropsychologia.2017.04.017.

Truskauskaitė-Kunevičienė, I. (2016). Direct and mediated long-term effects of the positive youth development intervention program try volunteering on empathy and prosocial behavior. *Social Inquiry into Well-being*, *2*(2), 8.

Urbina, D. A., & Ruiz-Villaverde, A. (2019). A critical review of homo economicus from five approaches. *American Journal of Economics and Sociology*, *78*(1), 63–93.

van de Groep, S., Meuwese, R., Zanolie, K., Guroglu, B., & Crone, E. A. (2018). Developmental changes and individual differences in trust and reciprocity in adolescence. *Journal of Research on Adolescence*, *30*(1), 192–208. https://doi.org/10.1111/jora.12459.

van de Groep, S., Zanolie, K., Burke, S. M., Brandner, P., Fuligni, A. J., & Crone, E. A. (2022). Growing in generosity? The effects of giving magnitude, target, and audience on the neural signature of giving in adolescence. *Developmental Cognitive Neuroscience*, *54*, 101084. https://doi.org/10.1016/j.dcn.2022.101084.

van de Groep, S., Zanolie, K., & Crone, E. A. (2020a). Familiarity and audience effects on giving: A functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience*, *32*(8), 1577–1589. https://doi.org/10.1162/jocn_a_01568.

van de Groep, S., Zanolie, K., & Crone, E. A. (2020b). Giving to friends, classmates, and strangers in adolescence. *Journal of Research on Adolescence*, *30*(Suppl. 2), 290–297. https://doi.org/10.1111/jora.12491.

van de Groep, S., Zanolie, K., Green, K. H., Sweijen, S. W., & Crone, E. A. (2020). A daily diary study on adolescents' mood, empathy, and prosocial behavior during the COVID-19 pandemic. *PLoS One*, *15*(10), e0240349. https://doi.org/10.1371/journal.pone.0240349.

van den Bos, W., van Dijk, E., Westenberg, M., Rombouts, S. A., & Crone, E. A. (2009). What motivates repayment? Neural correlates of reciprocity in the trust game. *Social Cognitive and Affective Neuroscience*, *4*(3), 294–304. https://doi.org/10.1093/scan/nsp009.

van den Bos, W., van Dijk, E., Westenberg, M., Rombouts, S. A., & Crone, E. A. (2011). Changing brains, changing perspectives: The neurocognitive development of reciprocity. *Psychological Science*, *22*(1), 60–70. https://doi.org/10.1177/0956797610391102.

van den Bos, W., Westenberg, M., Van Dijk, E., & Crone, E. A. (2010). Development of trust and reciprocity in adolescence. *Cognitive Development*, *25*, 90–102.

Van der Graaff, J., Carlo, G., Crocetti, E., Koot, H. M., & Branje, S. (2018). Prosocial behavior in adolescence: Gender differences in development and links with empathy. *Journal of Youth and Adolescence*, *47*(5), 1086–1099. https://doi.org/10.1007/s10964-017-0786-1.

van der Meulen, M., Steinbeis, N., Achterberg, M., Bilo, E., van den Bulk, B. G., van IJzendoorn, M. H., et al. (2017). The neural correlates of dealing with social exclusion in childhood. *Neuropsychologia*, *103*, 29–37. https://doi.org/10.1016/j.neuropsychologia.2017.07.008.

van der Meulen, M., Steinbeis, N., Achterberg, M., van IJzendoorn, M. H., & Crone, E. A. (2018). Heritability of neural reactions to social exclusion and prosocial compensation in middle childhood. *Developmental Cognitive Neuroscience*, *34*, 42–52. https://doi.org/10.1016/j.dcn.2018.05.010.

van der Meulen, M., van IJzendoorn, M. H., & Crone, E. A. (2016). Neural correlates of prosocial behavior: Compensating social exclusion in a four-player Cyberball game. *PLoS One*, *11*(7), e0159045. https://doi.org/10.1371/journal.pone.0159045.

Van der Meulen, M., Wierenga, L. M., Achterberg, M., Drenth, N., Van IJzendoorn, M. H., & Crone, E. A. (2020). Genetic and environmental influences on structure of the social brain in childhood. *Developmental Cognitive Neuroscience*, *44*, 100782. https://doi.org/10.1016/j.dcn.2020.100782.

van Goethem, A. A. J., van Hoof, A., van Aken, M. A. G., Orobio de Castro, B., & Raaijmakers, Q. A. W. (2014). Socialising adolescent volunteering: How important are parents and friends? Age dependent effects of parents and friends on adolescents' volunteering behaviours. *Journal of Applied Developmental Psychology*, *35*(2), 94–101. https://doi.org/10.1016/j.appdev.2013.12.003.

Van Hoorn, J., Van Dijk, E., Guroglu, B., & Crone, E. A. (2016). Neural correlates of prosocial peer influence on public goods game donations during adolescence. *Social Cognitive and Affective Neuroscience*, *11*(6), 923–933. https://doi.org/10.1093/scan/nsw013.

Van Hoorn, J., Van Dijk, E., Meuwese, R., Rieffe, C., & Crone, E. A. (2016). Peer influence on prosocial behavior in adolescence. *Journal of Research on Adolescence*, *26*(1), 90–100.
Van Leijenhorst, L., Gunther Moor, B., Op de Macks, Z. A., Rombouts, S. A., Westenberg, P. M., & Crone, E. A. (2010). Adolescent risky decision-making: Neurocognitive development of reward and control regions. *NeuroImage*, *51*(1), 345–355. https://doi.org/10.1016/j.neuroimage.2010.02.038.
Van Leijenhorst, L., Zanolie, K., Van Meel, C. S., Westenberg, P. M., Rombouts, S. A., & Crone, E. A. (2010). What motivates the adolescent? Brain regions mediating reward sensitivity across adolescence. *Cerebral Cortex*, *20*(1), 61–69. https://doi.org/10.1093/cercor/bhp078.
Veenstra, R., & Laninga-Wijnen, L. (2021). Peer network studies and interventions in adolescence. *Current Opinion in Psychology*, *44*, 157–163. https://doi.org/10.1016/j.copsyc.2021.09.015.
Vrijhof, C. I., van den Bulk, B. G., Overgaauw, S., Lelieveld, G. J., Engels, R. C., & van IJzendoorn, M. H. (2016). The prosocial Cyberball game: Compensating for social exclusion and its associations with empathic concern and bullying in adolescents. *Journal of Adolescence*, *52*, 27–36. https://doi.org/10.1016/j.adolescence.2016.07.005.
Westhoff, B., Blankenstein, N. E., Schreuders, E., Crone, E. A., & van Duijvenvoorde, A. C. K. (2021). Increased ventromedial prefrontal cortex activity in adolescence benefits prosocial reinforcement learning. *Developmental Cognitive Neuroscience*, *52*, 101018. https://doi.org/10.1016/j.dcn.2021.101018.
Westhoff, B., Molleman, L., Viding, E., van den Bos, W., & van Duijvenvoorde, A. C. K. (2020). Developmental asymmetries in learning to adjust to cooperative and uncooperative environments. *Scientific Reports*, *10*(1), 21761. https://doi.org/10.1038/s41598-020-78546-1.
Wierenga, L. M., Bos, M. G. N., Schreuders, E., Vd Kamp, F., Peper, J. S., Tamnes, C. K., et al. (2018). Unraveling age, puberty and testosterone effects on subcortical brain development across adolescence. *Psychoneuroendocrinology*, *91*, 105–114. https://doi.org/10.1016/j.psyneuen.2018.02.034.
Wray-Lake, L., DeHaan, C. R., Shubert, J., & Ryan, R. M. (2019). Examining links from civic engagement to daily well-being from a self-determination theory perspective. *The Journal of Positive Psychology*, *14*(2), 166–177. https://doi.org/10.1080/17439760.2017.1388432.
Yeager, D. S., Dahl, R. E., & Dweck, C. S. (2018). Why interventions to influence adolescent behavior often fail but could succeed. *Perspectives on Psychological Science*, *13*(1), 101–122. https://doi.org/10.1177/1745691617722620.
Yeager, D. S., Fong, C. J., Lee, H. Y., & Espelage, D. L. (2015). Declines in efficacy of anti-bullying programs among older adolescents: Theory and a three-level meta-analysis. *Journal of Applied Developmental Psychology*, *37*, 36–51. https://doi.org/10.1016/j.appdev.2014.11.005.

CHAPTER SEVEN

Gaze following in infancy: Five big questions that the field should answer

Kim Astor* **and Gustaf Gredebäck**
Uppsala Child & Baby Lab, Uppsala University, Uppsala, Sweden
*Corresponding author: e-mail address: kim.astor@psyk.uu.se

Contents

Abstract

Though much is known about the emergence and development of gaze following in infancy, there are large disagreements in some critical areas and major uncertainties within others. In this work, we highlight some of these areas in terms of five big questions that we believe are essential to address in order to advance research in the field. (1) How does social environment and culture impact gaze following? (2) What mechanisms drive the emergence of gaze following? (3) Does gaze following facilitate language development? (4) Is diminished gaze following an early marker of Autism? (5) How does gaze following relate to perspective-taking? This chapter aims not to answer these questions but to stimulate a discussion about the fundamental principles and assumptions on which the field resides and potentially serve as a guide for future research programs.

The ability to share attention is a central pillar in human social development (Emery, 2000; Mundy, 2016). It allows us to coordinate our behavior to, and learn from the behaviors of others with incredible efficiency and

Advances in Child Development and Behavior, Volume 63
ISSN 0065-2407
https://doi.org/10.1016/bs.acdb.2022.04.003

precision (Baron-Cohen, 1995; Zuberbühler, 2008). It is hard to imagine how civilization would be possible without this fundamental skill. However, it is not only humans that master this powerful tool. Attention sharing has a broad phylogenetic base, including a wide range of vertebrates such as non-human primates (Bräuer, Call, & Tomasello, 2005; Itakura, 1996; Tomasello, Call, & Hare, 1998; Tomasello, Hare, Lehmann, & Call, 2007), goats (Kaminski, Riedel, Call, & Tomasello, 2005), ravens (Schloegl, Kotrschal, & Bugnyar, 2007), and canines (Téglás, Gergely, Kupán, Miklósi, & Topál, 2012). In the animal kingdom, it is discussed as a process that allows detection of predators, food, or other relevant, often social, information. In humans, it is often regarded as a process that allows children and adults to read others' minds, and it is regarded as an efficient mechanism that infants can use to learn about their environment (Brooks & Meltzoff, 2015; Csibra & Gergely, 2009; Gredebäck, Astor, & Fawcett, 2018; Shepherd, 2010).

In this chapter, we will focus on the early development of the ability to share visual attention, or as we commonly refer to it, the ability to follow gaze. Though we, on the one hand, believe that attention sharing is critical to humanity, the specific role of gaze following appears less certain when put under the microscope. Despite a steady stream of publications on the development of gaze following, we would like to argue that scientific progress within the field is slow—that data is produced with a high frequency but without adding substantial advancements to our understanding of the phenomenon under investigation and its long-term development. We will highlight five big questions that we believe the field currently is unable to answer, either due to sparse data or a heavily divided research community. There are clearly other areas and ways to organize the future directions for research in the field of gaze following. But from our standpoint, these five questions capture some essential gaps in the literature, gaps that we need to fill in order to move the field forward. We see this as a first attempt to identify questions that, when answered, can alter our understanding and potentially create a paradigm shift for the field, with new understandings that potentially will reshape our understanding of gaze following.

The five big questions for the gaze following literature discussed in this paper are: (1) How does social environment and culture impact gaze following? (2) What mechanisms drive the emergence of gaze following? (3) Does gaze following facilitate language development? (4) Is diminished gaze following an early marker of Autism? (5) How does gaze following relate to perspective-taking? Before diving into these big questions, we will start by defining gaze following, describe the paradigms used to assess these abilities, and describe what we do know about its ontology in human infants.

1. Gaze following

1.1 Definitions

Responding to triadic joint visual attention describes in technical terms what we otherwise refer to as gaze following. Joint attention (JA) is a broad term that captures how individuals synchronize attention to the same foci point. One could either initiate joint attention (IJA), for example by pointing at something, or responding to a joint attention bid (RJA), a pointing gesture or, in this case, gaze direction. Though attention can be mediated by any modality, gaze following is specifically concerned with vision. Finally, joint visual attention can be established between individuals through dyadic eye contact, but in gaze following, it is also directed towards an external object; it is triadic.

1.2 The gaze following test paradigm

In the standard gaze following paradigm (designed to measure spontaneous gaze following), an interaction partner (actor) typically looks at the participant before shifting her gaze towards a target object. In infant studies, there are generally two objects placed on either side of the actor (target and distractor) within infants' visual field Fig. 1. Variations of this paradigm do exist where there, for example, are no gaze targets (Thorup, Nyström, Bölte, & Falck-Ytter, 2021), where there are more than two targets (von Hofsten, Dahlström, & Fredriksson, 2005), and where the targets are placed outside of infants' visual field (Deák, Flom, & Pick, 2000), but these are rare.

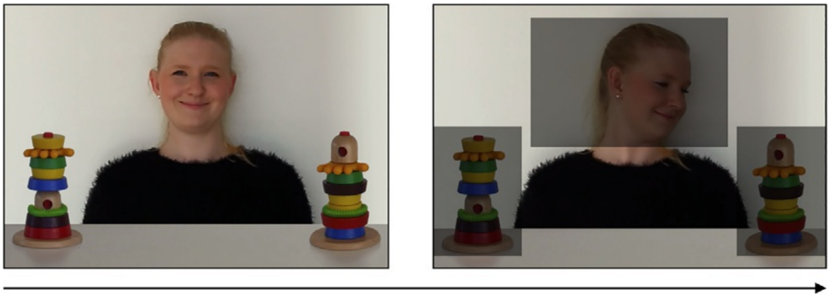


Fig. 1 An exemplar gaze following paradigm from Astor and Gredebäck (2019) utilizing eye-tracking technology. The left image illustrates the greeting phase where the actor establishes eye contact with the infant. Then, in the gaze phase (right image), the actor turns to look at an object at her side. Included are three areas of interest (black rectangles, not visible in the stimuli) used to calculate gaze following.

The response of following the actor's gaze direction is often seen as a deliberate, top-down action initiated by the infant (for alternative views, see Deák, 2015; Moore, 2008). Studies use different operational definitions to capture gaze following, depending in part on the technique used to record infants gaze following (e.g., video coding and eye-tracking): (1) accuracy of first look from the agent's face to a target (Astor, Thiele, & Gredebäck, 2021), (2) relative looking time to the target and distractor (Senju et al., 2015), (3) frequency/number of gaze alterations between gaze target and the agent (Hernik & Broesch, 2019), and (4) latencies to fixate the target (Gredebäck, Fikke, & Melinder, 2010). The first look metric is generally regarded as the gold standard gaze following metric, used in most studies, representing the accuracy of the initial response to another person's gaze shift (Gredebäck, Johnson, & von Hofsten, 2010).

2. Ontogeny

In their foundational study from almost 50 years ago, Scaife and Bruner (1975) found indications of visual co-orientation, later referred to as gaze following, in young infants. Though their study had methodological limitations (for an elaboration of these limitations, see Deák, 2015), their notion was correct, and later research has replicated their findings many times. Today we know that from about the age of 3 to 4 months, infants start to develop their joint attention skills beyond dyadic interactions, such as eye contact, to triadic interactions, such as gaze following (Astor et al., 2021; Astor & Gredebäck, 2019; D'Entremont, 2000; D'Entremont, Hains, & Muir, 1997; Gredebäck, Fikke, & Melinder, 2010; Gredebäck, Johnson, & von Hofsten, 2010; Michel, Kayhan, Pauen, & Hoehl, 2021). Though infants' gaze following ability is functionally limited by the time of emergence, it gets refined both quantitatively and qualitatively throughout infancy (Del Bianco, Falck-Ytter, Thorup, & Gredebäck, 2019).

Developmental progression during the first year after birth is demonstrated by the increased precision with which infants can track targets from other peoples' gaze. At 6 months of age, other objects that are presented along the scanning path (from an actor's face to the object that the actor is attending) distract infants' attention. Once infants disengage attention from the interaction partner, they will simply stop scanning to focus on the first salient object they encounter. At 12 months, infants start to follow gaze with greater precision, ignoring distractor objects along the scanning path to correctly distinguish the correct gaze target and follow gaze to targets outside the infant's visual field. At 18 months, infants follow others' line of regard to look at targets located behind themselves (Butterworth & Cochran, 1980; Butterworth & Jarrett, 1991; Deák et al., 2000; Morissette, Ricard, & Gouin Décarie, 1995). Other studies find that infants become faster to respond to gaze following bids the older they get (Gredebäck, Fikke, & Melinder, 2010; Gredebäck, Johnson, & von Hofsten, 2010) and that they rely more and more on social rather than perceptual information, when presented together, during their first year (Astor et al., 2021).

Infants' tendency to follow gaze is impacted by their arousal and the attentional state (Gredebäck et al., 2018; Ishikawa, Yoshimura, Sato, & Itakura, 2019; Szufnarowska, Rohlfing, Fawcett, & Gredebäck, 2014) with more gaze following if the actor makes eye contact, addresses the infant, or acts in a novel and surprising manner before shifting gaze to a target. In addition, the more salient the gaze shift cue is, the more likely are infants to follow gaze to this object (Deák et al., 2000). Though motion seems to be more important at the age of onset (Astor et al., 2021; though other low-level cues such as proximity do not evoke similar effects; Astor & Gredebäck, 2019), later, at the age of 9 months, infants who already follow gaze spontaneously do not need to see the motion (Moore, Angelopoulos, & Bennett, 1997). Sensitivity to salience also applies to gaze targets. Infants demonstrate more gaze following in the presence of moving targets (Butterworth, 1991), and they follow gaze to targets more often if the target is higher in complexity and novel in the context of distractor objects (Deák et al., 2000). Finally, when infants start to follow gaze, they tend to follow an interaction partner's head direction (Michel et al., 2021), but over the following months, the importance of the interaction partner's eyes increases (Brooks & Meltzoff, 2002, 2005; Meltzoff & Brooks, 2007, 2008; Tomasello et al., 2007).

3. Five big questions

Much progress has been made in the endeavor to understand the development of gaze following in infancy, and there is, according to our reading of the literature, a reasonable consensus about the developmental trajectory of gaze following behavior early in life (outlined briefly above and reviewed by Del Bianco et al., 2019). However, despite an active research community and a steady stream of published research, there are large disagreements in some critical areas and major uncertainties within others, areas that go beyond mapping out the development of gaze following on a behavioral level. In this chapter, we highlight five big questions that we believe are essential to address in order to truly understand gaze following, both as a phenomenon on a mechanistic and theoretical level and regarding its role in early development more broadly. These questions focus on assumptions made about the purpose, function, and/or variability of gaze following in infancy and early childhood. These five big questions are: (1) How does social environment and culture impact gaze following? (2) What mechanisms drive the emergence of gaze following? (3) Does gaze following facilitate language development? (4) Is diminished gaze following an early marker of Autism? (5) How does gaze following relate to perspective-taking?

4. Question I: How does social environment and culture impact gaze following?

Despite active ongoing research on the development of gaze following, there is one area that has received very little attention: What is the impact of the environment in general and culture in particular? A few studies demonstrate that gaze following tendencies are influenced by infants' early experiences. Astor et al. (2020) found that infants who demonstrated a secure attachment style followed gaze more at 6 months of age, and that infants of mothers with elevated depressive symptoms instead followed gaze to a lesser extent at 10 months of age. To our knowledge, this is the first attempt to address the link between the quality of infants' social and emotional context and gaze following. Research focusing on preterm infants provides some information regarding the quantity of environmental experience, suggesting that the amount of experience rather than biological maturation (postmenstrual age) determines the tendency to follow gaze (De Schuymer, De Groote, Striano, Stahl, & Roeyers, 2011; Peña, Arias, & Dehaene-Lambertz, 2014).

Senju et al. (2015) investigated the consequences of being raised by blind compared to seeing parents. In contrast to studies on premature infants, their results indicate that the amount of experience[a] during the first year after birth

[a] Note that this is different from the experience in studies of premature infants. Instead of overall life experience, the experience targeted in studies of infants with blind parents is that interaction with a partner that reliably indicates interesting objects with gaze. The assumption is that blind parents' "gaze direction"less frequently points at interesting things, making it less rewarding to follow, and resulting in less gaze following.

does not impact the emergence of gaze following. After their first birthday, however, infants of blind parents demonstrated less looking time towards gaze targets (indicating less referential expectation), though their initial orientation to the target was still equally frequent as infants of seeing parents. Brooks, Singleton, and Meltzoff (2020) investigated gaze following in deaf infants of deaf parents. They instead found enhanced gaze following in these infants, compared to hearing infants.

In 2019, Hernik and Broesch published results from rural Vanuatu (South Pacific Ocean), where face-to-face interactions are less frequent than in Western countries. Using three metrics of gaze following, they found that 5- to 7-month-old Vanuatu infants looked longer (indicating referential understanding) and more frequently at the gaze target. However, as indicated by their first gaze shift, infants did not make an initial accurate reorientation to align their gaze with the actor. In a study with adult participants, Cohen, Sasaki, German, and Kim (2017) found that adults rely more on group consensus when following gaze in more interdependent cultures (east Asia). European Americans, in contrast, largely ignored the group and followed the leader's gaze (foreground person vs background individuals). Based on the limited data we have access to, it seems that gaze following is influenced by both quantity and quality of social context and experience.

Infant-parent social interactions - the major part of infants' social environment - varies substantially across cultures (Keller, 2007). For example, there is an emphasis on face-to-face interactions, cognitive development, and agency in Western cultures. By contrast, physical contact, motor development, and social integration tend to be valued more highly in majority world contexts (Bornstein, 2013; Keller, 2007). There are also cultural variations related to family structure, with more extended families in majority world contexts (Kağitçibaşi, 2007). Family bonds and secure attachment relations are expressed differently across cultural contexts (Keller, 2018). There are also pronounced differences in the amount of non-verbal communication (Nsamenang & Lamb, 2014) and parental views on infants' basic capabilities (Kotchabhakdi, Winichagoon, Smitasiri, Dhanamitta, & Valyasevi, 1987) across cultures.

All of these cultural differences are well documented, but these differences (that also exist across families within a cultural context) are not taken into account when gaze following is studied. There is enough research to suggest that gaze following is an experience dependent process and that qualitative and quantitative variability across cultural contexts impacts the development of gaze following in some way. Unfortunately, the overwhelming majority of studies on gaze following are conducted in western urban settings, making it difficult to generalize knowledge across contexts.

Furthermore, and perhaps most importantly, for long, many developmental psychologists have taken for granted that findings in western settings generalize across cultures (for an elaboration on this issue, see Keller, 2018). This also becomes evident in the gaze following literature as the vast majority of studies ignore this question while simultaneously constructing theoretical frameworks that claim universality (Astor et al., 2020; Csibra & Gergely, 2009; Meltzoff, 2007; Triesch, Teuscher, Deák, & Carlson, 2006). Investigating the influence of social, emotional, and cultural contexts is critical if we want to create a more inclusive model of gaze following and empirically test assumptions of universality. Current data suggest that social and emotional environment and culture impact early developing gaze following, but we have a long way to go before we as a field understand how these factors interact in order to shape the development of gaze following.

5. Question II: What mechanisms drive the emergence of gaze following?

One question that dates back to Scaife and Bruner's (1975) original study concerns the mechanisms of emerging gaze following. Scaife and Bruner speculated that infants' ability to coordinate visual attention with others might be a sign of social awareness, indicating that infants were less than completely egocentric (a common notion at the time; Piaget, 1952, though see Kesselring & Müller, 2011 for an elaboration of Piaget's notion of egocentrism). Almost 50 years later, the field still largely disagrees on this point, and the mechanisms governing emerging gaze following is probably the most controversial question in the gaze following literature today (Del Bianco et al., 2019). At the same time, the status of gaze following as an early emerging ability, with significance for later development, makes it essential to understand the principles of its development. Though there is a wide range of theories trying to explain different stages of development, three broad theoretical themes on the developmental origin of gaze following can be distinguished. First, a nativist perspective suggests that emerging gaze following is domain-specific and driven by social awareness. In contrast, reductionist views suggest that domain-general perceptual-motor responses guide infants' visual attention in social interactions. Finally, an empiricist perspective suggests that infants acquire gaze following capability through a domain-general learning process.

There is a wide agreement that gaze following, at some point, does reflect social awareness. We refer to theories that interpret infants' early manifestations as corresponding to social information processing as the social awareness perspective. This nativist perspective consists of a heterogeneous group of theories that suggest different mechanisms as the foundation of social gaze following. Unifying these theories is the assumption that a domain-specific social process drives this behavior. For example, it has been suggested that infants are born with specialized neural processing modules that allow them to share attention and encode eye direction in a reflexive manner (Baron-Cohen, 1995). He suggested that Fodorian modules, such as an eye direction detector, process specific categories of sensory input. Support of this notion is found in research indicating that newborns prefer looking at faces with open eyes (Batki, Baron-Cohen, Wheelwright, Connellan, & Ahluwalia, 2000) and at faces displaying direct over averted gaze (Farroni, Csibra, Simion, & Johnson, 2002; Farroni, Massaccesi, Pividori, & Johnson, 2004). Other theories are more process-oriented. Meltzoff (2007) proposed the idea that infants recognize others as similar to themselves (the "like me" hypothesis) and therefore understand that if a person is looking in any given direction, this corresponds to what they would experience if they performed a similar action. Natural pedagogy (Csibra & Gergely, 2009) takes a somewhat different approach and suggests that infants' gaze following is specific to contexts where they expect to learn something from the interaction partner. Infants' expectation to learn is believed to be conditioned by perceived infant-directed communication. From this perspective, infants are thought to have an innate sensitivity to so-called ostensive (communicative) cues. In support of this notion, studies have reported more gaze following if the gaze shift is preceded by an ostensive (communicative) cue, such as direct gaze, contingent interaction, or infant-directed speech (e.g., Hernik & Broesch, 2019). A final theory focuses on attentional mechanisms and infants' motivation to engage with others (Astor et al., 2020; Gredebäck et al., 2018; Gredebäck, Fikke, & Melinder, 2010; Gredebäck, Johnson, & von Hofsten, 2010). Note that social awareness theories vary in how they look at the role of infants' environment. While Natural pedagogy and the Fodorian modularity theory point at mechanisms that suggest innate experience-expectant preparedness, the "Like me" hypothesis and the theory of attention and social motivation suggest an experience-dependent process (see Fig. 2).

	Experience-dependent	***Experience-expectant***
Social/ Domain-specific	Social motivation "Like me"	Fodorian modularity Natural Pedagogy
Non-social/ Domain-general	Reinforcement	Prepotent rapid learning Perceptual cueing

Fig. 2 The X-axis corresponds to environmental influence, and the Y-axis corresponds to domain specificity (or social awareness).

Perceptual cueing is a perspective that suggests that emerging gaze following in young infants does not require higher-level cognition such as social motivation. Instead, it suggests that it is indistinguishable from domain-general perceptual-motor development. The foundation for the perceptual cueing perspective thus relies on a reductionistic, Occam's razor, approach, suggesting that we should remove unnecessary complexity from a model in order to reach a stronger explanation. In support of this perspective is studies demonstrating that infants are bounded by their perceptual capabilities and influenced by bottom-up responses. For instance, they only search for objects within their own visual field, and further, they scan in the general—rather than the specific—direction of the cue (Butterworth & Cochran, 1980; Butterworth & Jarrett, 1991), seemingly unable to comprehend that other's point of reference is different from their own. Indeed, subsequent research has shown that infants at the age of 9 months are cued by gaze only if they observe the motion of the head turn (Moore et al., 1997).

Similarly, it has been demonstrated in 4- to 5-month-olds that the salient motion of an observed head turn overrides relative eye displacement, and further, that infants redirect their gaze faster and more frequently in the same direction as the pupils of a face only if they observed the motion of the pupil (Farroni, Johnson, Brockbank, & Simion, 2000). Interestingly, they are cued by eye motion even when the eye moves from an averted position to direct gaze (Farroni, Mansfield, Lai, & Johnson, 2003). Later, Farroni et al. (2004) demonstrated that even gaze-cueing in newborns is dependent on perceived motion. In addition, studies using non-human stimuli such as abstract 3D objects (Deligianni, Senju, Gergely, & Csibra, 2011) or highly stylized schematic face-like stimuli (Farroni et al., 2004) suggest that the encoding of directional motion in similar paradigms adhere to a general process rather than being specific to the interactions with other conspecifics. A lot of work has focused on the competition between head and eye cues, and it is now well known that infants at an initial state rely more on the head as a general cue of direction. This suggests that it is not gaze comprehension but the perceived motion that guides infants' attention. Note that a substantial amount of studies taken to support the perceptual cueing perspective comes from the gaze cueing (c.f., gaze following) literature, reflecting the view of gaze following as a reflective, bottom-up process (see Fig. 2).

In the mid '90s, Corkum and Moore (1998) and Moore and Corkum (1994) offered an alternative, empiricist view on gaze following emergence. They suggested that gaze following is the result of basic hedonistic principles and a domain general learning process. In a series of experiments, Corkum and Moore demonstrated that 8-month olds who do not follow gaze spontaneously can learn this through reinforcement. Supporting this view is research demonstrating 14-month-olds ability to assess the cue validity of gaze on an individual basis (Chow, Poulin-Dubois, & Lewis, 2008). The reinforcement perspective has gained a lot of traction since it was first introduced and is now a prominent theory of gaze following emergence (e.g., Deák, Krasno, Triesch, Lewis, & Sepeta, 2014; Silverstein, Feng, Westermann, Parise, & Twomey, 2021). In this context it should be noted that a reinforcement model of emerging gaze following does not have to be a model absent of social development/skills/perception. In fact, the opposite is perhaps a more accurate description of this view. For example, Triesch and colleagues (Fasel, Deák, Triesch, & Movellan, 2002; Triesch et al., 2006) have suggested a minimum requirement model, a "basic set" of reinforcement learning mechanisms for gaze following emergence that includes a social bias.

A common notion within the reinforcement perspective is that infants' existing tendency to orient in the same direction as others (following principles of perceptual cueing) is refined by reinforcement (e.g., by rewarding sights at the target location). As such, gaze following is thought to develop in an invariant three-step process that incorporates components (social and perceptual information) from the other two frameworks: Very young infants respond in a reflexive manner to perceptual cues such as perceived motion and orient in the same direction (Deák, 2015; Moore & Corkum, 1994). Based on this rudimentary ability, infants fine-tune the association between gaze cues and "interesting sights" through a reinforcement learning process (Corkum & Moore, 1998). Co-occurrence of a perceived behavior (an interaction partner orients their gaze to an object), a behavior of the infant (to look in the direction being cued by gaze), and a reward (an interesting object or event appear in the line of sight) strengthens the tendency to align gaze with others. This process shapes infants' behavior over time and continues to fine-tune the gaze following response. Once this behavioral pattern is established, infants' may learn to appreciate the importance of gaze and begin to understand its social meaning. That is, as a final step in this process, infants/children bootstrap a social cognitive framework to a gaze following skill that is already quite sophisticated (Deák, 2015; Moore, 2008; Moore & Corkum, 1994).

A tentative reinforcement process is often conceptualized as a slow experience dependent process that shapes infants' tendency to follow gaze throughout development. However, a reinforcement process does not necessarily have to be slow, as noted in previous work (Astor et al., 2020). In fact, studies demonstrating relatively fast learning that seem to rely on both saliency principles (Moore et al., 1997) and ecological validity as Corkum and Moore (1998) were unsuccessful in training infants to misalign with the gaze direction (also see Michel et al., 2021). These studies can be argued to suggest a fast experience-expectant process where gaze following might emerge based on rapid learning of a prepotent stimulus-response (potentially putting it on the border between Social—domain-specific and non-social domain-general processes, see Fig. 2).

Three broad theoretical perspectives on the developmental origin of gaze following, rooted in nativist, reductionist, and empiricist traditions, have been covered. In an attempt to arrange specific theories within these perspectives, we suggest that they can be divided along two axes, one that addresses social awareness and one that concerns environmental influence on development (see Fig. 2). Though there are many different ways to conceptualize and arrange theories, recent work has demonstrated that this might be a productive approach (Astor et al., 2020).

In contrast to active theoretical discussions, real theoretical advancements, that unify the field, are absent. Much research is confirmatory in nature, and surprisingly few studies try to explicitly test the theories against each other. However, only putting the conventional frameworks against each other might not be the solution that advances our understanding. While theories often appear to make exclusivity claims, it is also possible that gaze following develops across multiple paths. In fact, many authors have argued that the underlying mechanisms of gaze following are likely to undergo change during development (Brooks & Meltzoff, 2005; Deák, 2015; Del Bianco et al., 2019; Jasso, Triesch, Deák, & Lewis, 2012; Moore, 2008). Others have suggested that gaze following develops along parallel paths (Shepherd, 2010), a notion that is consistent with Astor et al. (2021), who report early sensitivity to both perceptual (motion) cues and gaze direction.

A possible way forward is large-scale collaborations were predictions from different theories are developed and examined together, not in isolated pipelines by individual research groups. A better theoretical understanding of the mechanisms involved in the development of gaze following would make it possible to explore how different mechanisms might interact during development in general and with regard to related abilities (e.g., language), clinical assessment (e.g., Autism), and environmental contexts (e.g., culture) in particular. For now, this question remains unanswered. The field is nowhere near formulating a robust process model accounting for the mechanisms behind gaze following emergence.

6. Question III: Does gaze following facilitate language development?

The view on gaze following as a facilitator of language development is widely accepted in the gaze following literature (Del Bianco et al., 2019; Mundy, 2016; Shepherd, 2010). Some suggest that gaze following is necessary for successful word learning (Gliga et al., 2012), and others that it is a critical facilitator of language acquisition (Dehaene, 2020). Let us illustrate with an example: A mother plays with her infant, alternating her gaze between the infant and the toys around them. During play, the mother says things like "Oh, a doll!", while at the same time looking at an object with four limbs and a cute face. Following the mother's gaze to look at the doll allow the infant to pick up on the word "doll" and map it to the object referred to as a doll. According to the gaze following literature, this is how infants use others' gaze to learn words. Unfortunately, there might be good reasons to question this view.

In a recent review on the role of gaze following on language development in infancy, Çetinçelik, Rowland, and Snijders (2021), conclude that there is substantial evidence for longitudinal relationships between gaze following in infancy and later language development. Their conclusion aligns with the consensus in the literature at large, a literature frequently citing this association to highlight the significance of gaze following in development. Whenever there is a general agreement on an issue that simultaneously serves to legitimize existing research programs, such as the one we observe here, it is worth scrutinizing. Some notable limitations in influential work motivated us to look further and ask if gaze following facilitates language development, a question that many, including ourselves, believed to have already been answered. Here we will take a closer look at the empirical support for the association between gaze following and language development.

There are a few lines of research concerned with the role of gaze following in language development. Here we will focus on correlational longitudinal work. In this context, one of the most influential articles to date was written by Brooks and Meltzoff (2005). In their study, levels of gaze following at the age of 10–11 months (combined age group) were used to predict language at 14 and 18 months. To measure language, they used four CDI (Communicative Development Inventory; Fenson et al., 1994) subscales at each age for a total of eight correlations. Only one subscale, "total gestures" at 18 months, correlated with earlier gaze following. Word production and comprehension did not correlate with gaze following at any time point. We also note that though 9-month-olds were tested, the results for this age group were not reported. In addition to the gaze following measure, this study measures infant vocalization as a potential predictor of language, though none of the eight correlations between vocalization and later language were significant. However, when they combined the two independent measures, gaze following and simultaneous vocalization, this composite score predicted language development in six out of eight instances. Specifically, language comprehension and total gestures, but not word production, were associated with the 14 and 18 months composite scores. Consistent with later literature citing this study, the authors describe their results in terms of a strong correlation between gaze following and subsequent language development. While Brooks and Meltzoff's results are exciting, they are complex, and from this study alone, the role of gaze following in language development is not clear.

Looking at another influential study, Morales et al. (2000a) reported infants' gaze following at 6 months and simultaneous gaze and point following at seven additional occasions, between 8 and 24 months, in relation to language development. Language ability was assessed at 24 months with one expressive measure and later, at 30 months, with two expressive and one receptive measure. At 6, 8, 10, 12, and 18 months, gaze/point following predicted at least one of four measures of later vocabulary development. At 15, 21, and 24 months, gaze and point following was not associated with any instance of vocabulary development. Interestingly, gaze following at 6 months seemed to be the best predictor, associated with expressive language at 24 and both expressive and receptive language at 30 months.

Perhaps the most cited study in this context (3468 citations, Google scholar 2022-01-06) was conducted by Carpenter, Nagell, Tomasello, Butterworth, and Moore (1998). They investigated a wide range of temporal associations between gaze and point following and language development.

They tested infants on a monthly basis, between 9 and 15 months, and found two significant associations, from gaze and point following at 9 and 10 months to language at 12 months. It is not clear how many cross-lagged correlations were performed in this study. If they allowed predictions both ways and co-occurring measures, that would be a total of 49 correlations. If language was not allowed to predict attention following backward in time, it is 28 correlations. The statistical power for these analyses appeared to be small. Some measures were particularly affected; these were, for example, instances where almost none or all infants had passed the test. However, when looking at the age of emergence, including all assessments, they did find a correlation indicating that the younger infants are when they start to respond to joint attention the earlier they express referential language.

Though these studies are frequently cited in the gaze following literature, they can inspire further discussions. While they find positive correlations to later language development, they either depend on other language-related variables, such as vocalizing, or the significant correlations are contrasted with a large number of null results. Another issue specific to the latter two of these studies is that they do not assess gaze following specifically but include it, together with point following, in a broader construct aiming to assess RJA. Looking at these three studies, it is difficult to evaluate the contribution of gaze following to language development, which motivates a further examination of the existing literature. We will rely on the work of Çetinçelik et al. (2021) to examine whether the studies included in their review do, in fact, provide a solid foundation for the current narrative; that gaze following facilitates language development.

Bellow we provide a graphical overview (Fig. 3) presenting results from studies included in Çetinçelik et al. (2021). In addition to the studies covered above, this overview also includes Beuker et al. (2013), Brooks and Meltzoff (2008), Brooks and Meltzoff (2015), De Schuymer et al. (2011), Markus et al. (2000), Mundy et al. (2007), Mundy and Gomes (1998), Mundy et al. (2003), Morales et al. (2000b), Morales et al. (1998), and Tenenbaum et al. (2015). If there were multiple instances of significance testing, we only report the primary measure. If zero-order correlations are available, results from these are reported rather than full models. We only include predictions from joint attention to language, even though some studies look at reversed associations. We also excluded cases with seemingly insufficient variance, judged to be too low to detect correlations. Finally, only studies and assessments that capture RJA measures were included.

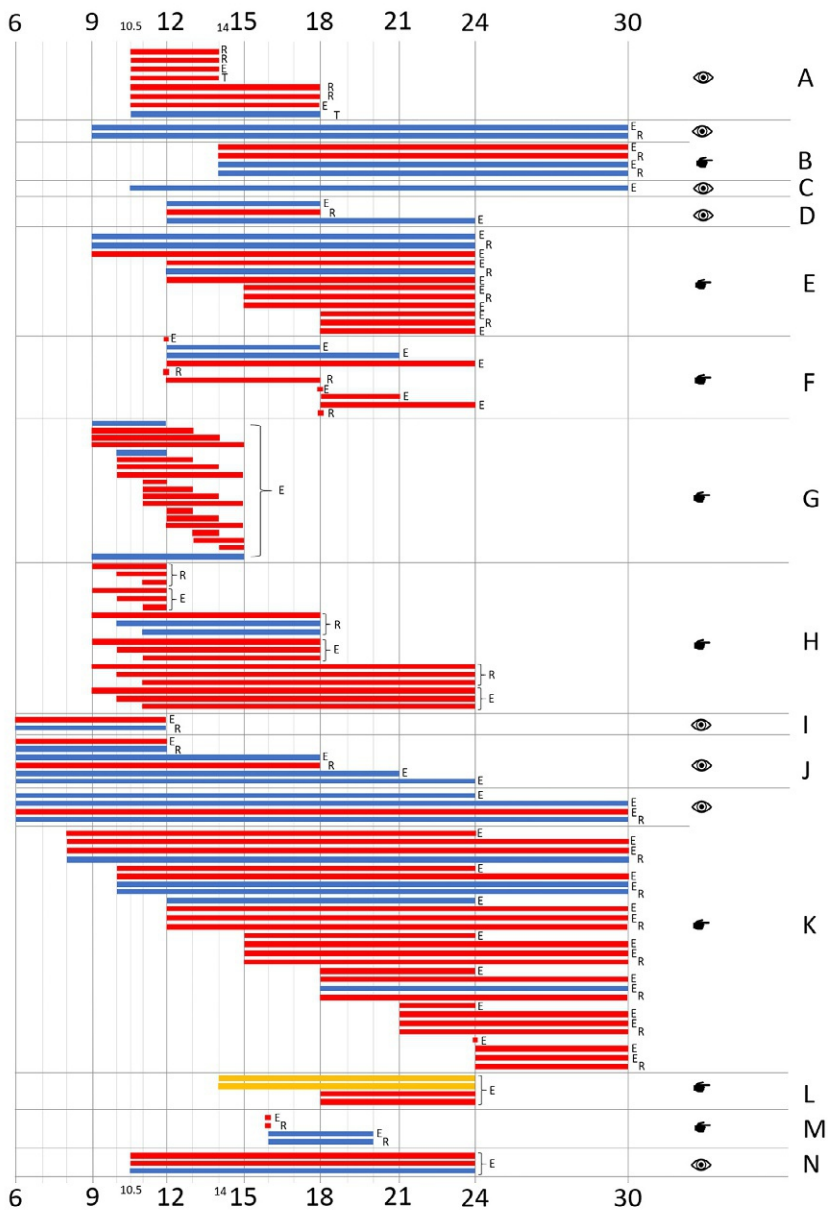


Fig. 3 Blue bars indicate significant positive associations. Red bars indicate non-significant associations. Yellow bars indicate one-tailed significant positive associations. Vertical lines with associated numbers correspond to the age of the infant. Each bar starts at the age for point and gaze following assessment and ends at the age for language assessment. The letter at the right end of the bar indicates expressive

This graphical overview did not include assessments combining RJA and IJA or those that capture IJA rather than RJA. As is evident from Fig. 3, the results appear inconclusive. Approximately one out of four tests demonstrate an association between RJA and later language, a number that would decrease further if all tests from these papers were included. Furthermore, as evident from Fig. 3, the timing of positive associations is not replicated in a systematic manner across studies. However, looking at the smaller number of studies concerned with gaze following specifically, almost half of the reported correlations are significant.

Aside from the correlational studies covered by Çetinçelik et al. (2021), there are other studies looking at the relation between gaze following and language, not included in their review. We will highlight a few of these studies here but not include them in the graphical overview since they were not retrieved systematically and may be prone to selection bias, but they are worth noting since they applied standard gaze following assessments (not including point following). For example, Senju et al. (2015), who investigated the development of gaze following comparing infants of blind and infants of sighted parents (see Question I), also assessed language. They found no relation between gaze following and receptive language in either group or visit (6–10 and 12–16 months). Juvrud et al. (2019) found an association between gaze following at 9 months and language development at 9 months, but not 18 or 24 months. In this study the CDI scale "words and gestures" was used to assess the 9-month-olds, and "words and sentences" for 18- and 24-month-olds. Here it seems possible that gestures drove the effect at the 9 moths' assessment, but again, similar to studies combining gaze and point following, it is impossible to identify the effect of a specific factor. In contrast, Slaughter and McConnell (2003) found a link between gaze following and productive language at 8 to 14 months, and

language (E), receptive language (R), and total gestures (T). The pointing hand symbol at the right side of the bars indicates that gaze and point following were combined to predict language. The eye symbol indicates standard gaze following assessments. Most studies include multiple assessments. Studies are separated by horizontal gray bars. Letter A to N correspond to the following studies: (A) Brooks and Meltzoff (2005), (B) De Schuymer et al. (2011), (C) Brooks and Meltzoff (2015), (D) Tenenbaum, Sobel, Sheinkopf, Malle, and Morgan (2015), (E) Mundy et al. (2007), (F) Markus, Mundy, Morales, Delgado, and Yale (2000), (G) Carpenter et al. (1998), (H) Beuker, Rommelse, Donders, and Buitelaar (2013), (I) Morales et al. (2000b), (J) Morales, Mundy, and Rojas (1998), (K) Morales et al. (2000a), (L) Mundy, Fox, and Card (2003), (M) Mundy and Gomes (1998), (N) Brooks and Meltzoff (2008).

finally, Okumura, Kanakogi, Kobayashi, and Itakura (2017)[b] found that both looking duration and first look predicted productive vocabulary at 18 months. This additional small sample of studies appears similarly inconclusive.

Looking beyond correlational longitudinal work, there have been experimental attempts to capture gaze following-mediated word-object mapping. Hollich et al. (2000) found that 24-month-olds could learn the name of an object using gaze cues, even in the presence of more salient distractor objects. However, 12-month olds only learned the name of novel objects if the interaction partner touched the object in addition to gaze cues. Barry-Anwar, Burris, Graf Estes, and Rivera (2017) demonstrated that infants (12 and 18 months) learned to map a label to an object if the interaction partner was their caregiver but not if it was a stranger. In contrast, Parsons et al. (2019) found no gaze following mediated word-object mapping in 15-month-olds. Once again, the results are inconclusive.

Today, there is a broad agreement in the gaze following literature that gaze following facilitates language development. However, when looking closer at studies exploring this hypothesis, it becomes clear that such claims are overconfident. There are a couple of issues worth paying attention to based on our review. First, and perhaps most striking, looking at the collection of results from the overview (Fig. 3), the majority are null results, showing no link between gaze following and language. It can be argued that some of the null results are expected because of a lack of statistical power. It can also be argued that measures of expressive language at young ages are doomed to fail (see Morales et al., 2000b). However, this is not enough to account for the low consistencies that we report.

Another critical issue is that a large number of these studies do not use standard gaze following measures. Instead, they combine gaze following with other RJA measures, often point following. This is especially noteworthy since there seems to be no clear association between point and gaze following (Carpenter et al., 1998; Carpenter, Pennington, & Rogers, 2002), and because gaze following and point following have an additive effect on infants' attention following, where pointing gestures appear more salient than gaze shifts (Deák et al., 2000). It is worth mentioning that composite scores (such as combining gaze and point following) are only a problem in

[b] Though this study was cited in Çetinçelik et al. (2021), it was not covered in their section concerned with correlations to later language development, and thus not part of the reviewd material used in their evaluation of that question.

contexts where we try to understand the role of specific abilities, such as the role of gaze following in language development. Clearly, many of the studies covered are not specifically targeting the connection between gaze following and language. Taken together, we believe that the role of gaze following in language development is still an open and important question.

7. Question IV: Is diminished gaze following an early marker of Autism?

Autism is a neurodevelopmental disorder characterized by difficulties in non-verbal communication and repetitive and restricted behavior (American Psychiatric Association, 2013), affecting about 1% of the population or more (Christensen et al., 2016; Nygren et al., 2012). Concerning the specific challenges in Autism, joint attention deficits are central to the difficulties with non-verbal communication (Filipek et al., 1999; Mundy, 2016), and the American Psychiatric Association (2013) describes the lack of gaze following in particular as a defining feature of Autism. Since joint attention skills such as gaze following emerge early in infancy, long before an Autism diagnosis is usually established (Ozonoff et al., 2015), gaze following has been recognized as a potential early marker of Autism (e.g., Baron-Cohen, 1995). Atypical gaze following in Autistic children has also been suggested to explain the difficulties observed in language development in this population (Baron-Cohen, Baldwin, & Crowson, 1997; Gliga et al., 2012). As a result, the link between gaze following and Autism is (similar to the link between gaze following and language; Question III) often highlighted in research that highlights the significance of gaze following during infancy (Astor & Gredebäck, 2019; Brooks & Meltzoff, 2005; Emery, 2000; Triesch et al., 2006; Zuberbühler, 2008). However, researchers targeting this issue have recently started to question this notion (Nyström, Thorup, Bölte, & Falck-Ytter, 2019; Parsons et al., 2019). Here we review a sample of papers investigating the association between Autism and gaze following, some of which frequently appear in the gaze following literature, fueling the notion of diminished gaze following as an early marker of Autism.

Though many studies investigate joint attention in Autistic populations, there is a limited amount of work targeting gaze following specifically. Of those that do, only a small number assess infants prospectively. Bedford et al. (2012) assessed infants' ability to follow gaze at 7 and 13 months. They found no difference in gaze following (first look) when comparing a group at-risk for Autism to a low-risk sample. Both groups followed gaze at both timepoints. However, 13-month-olds that received an Autism diagnosis at 36-month or were identified as having socio-communication difficulties, spent less attention (looking time) to the gaze target. This might indicate intact gaze following but challenges understanding the referential nature of others' gaze. The authors discuss the possibility that difficulties in social communication in general, not Autism specific, might underlie diminished gaze following in infancy.

In contrast to Bedford et al., Thorup, Nyström, Gredebäck, Bölte, and Falck-Ytter (2016) reported diminished gaze following in 10-month-old infants with familial risk of Autism. Diminished gaze following was specific to eye gaze, responding to head and eye gaze seemed intact. In a follow-up study (assessing the same sample as Thorup et al., 2016), Nyström et al. (2019) investigated whether differences found in the high-risk sample were indicative of the results in the group of infants who later received a diagnosis. This was not the case. Instead, infants with a later diagnosis demonstrated gaze following capacity similar to control infants. The same sample was re-examined by Thorup et al. (2021), who found indications of diminished referential expectation in infants with a later diagnosis. In the absence of gaze targets, these infants looked less at the gaze-referred area before looking back at the actor.

In a study with 15-month-olds, Parsons et al. (2019) found no difference in gaze following frequency when comparing infants with and without familial risk of Autism. Those who received an Autism diagnosis, however, spent less time looking at the objects and more at the face of the interaction partner, a pattern associated with later language skills and consistent with other findings (Bedford et al., 2012; Thorup et al., 2021). Though language deficiencies are no longer a criterion for Autism, it is common in this population. These results suggest a possible pathway for this association. Note that other visual attention factors related to Autism, such as "sticky attention," or delays in visual disengagement, evident at the age of 12 months (Elsabbagh et al., 2013; Sacrey, Bryson, & Zwaigenbaum, 2013), might help explain these dwell time discrepancies.

Findings in high-risk samples are sometimes discussed in terms of a broader autistic phenotype (BAP) that might not reflect clinical Autism per se, but rather Autistic-like tendencies, such as difficulties in social communication, and are more frequent in families of Autistic members. Landa, Holman, and Garrett-Mayer (2007) assessed infants' joint attention with tasks consisting of simultaneous pointing and gazing at targets at 14 and

24 months. At the 14-month assessment, infants who received an early diagnosis, at 14 months did not demonstrate less gaze following than the low risk Control or BAP group (in this study defined as high risk infants with language and/or social difficulties, as determined by standardized tests, but without an ASD diagnosis), but lower than non-BAP (infants who did not meet the criteria for ASD or BAP but who demonstrated motor or language difficulties not detected by the standardized tests). However, infants with an early diagnosis did perform worse than all other groups on the IJA task. Infants with a later diagnosis (at 24 months of age) did not perform differently than any other group at that time. At the 24-month assessment, both infants with early and late diagnosis performed worse than BAP, non-BAP, and Control groups on both RJA and IJA. Rozga et al. (2011) found that infants with a later Autism diagnosis did not demonstrate RJA (in response to simultaneous point and gaze) or IJA to the same degree as the typically developing control group at 12 months. High risk infants without Autism performed similar to the Control group on both IJA and RJA.

The current narrative in the gaze following literature implies that gaze following is affected in infants later diagnosed with Autism. This has become a common and effective way to highlight the significance of gaze following early in life. However, this assumption rests on a weak foundation as prospective infant studies looking specifically at gaze following do not find group differences based on clinical outcome. Though non-verbal social communication is affected in Autism (by definition APA, 2013), it is less clear if infants who receive a later diagnosis have specific difficulties using or understanding referential eye gaze in communication early in life, though some results point in this direction (Bedford et al., 2012; Parsons et al., 2019; Thorup et al., 2021). In contrast, studies on older children have demonstrated lower levels of gaze following in Autistic children (Carpenter et al., 2002; Gillespie-Lynch, Elias, Escudero, Hutman, & Johnson, 2013; Leekam, Baron-Cohen, Perrett, Milders, & Brown, 1997; Leekam, Hunnisett, & Moore, 1998).

If gaze following starts intact but later deviates from typical development, this could have many possible explanations, including developmental regression (Ozonoff et al., 2010) and a change in the mechanisms driving gaze following through development (see Astor et al., 2021). Researchers have argued that the underlying mechanisms of gaze following are likely to undergo change during development before it can be accurately described as a social ability (Brooks & Meltzoff, 2005; Deák, 2015; Moore, 2008) or develops across parallel paths utilizing different mechanisms simultaneously (e.g., reflexive responses and cognitive comprehension) (Shepherd, 2010). Astor et al. (2021) recently suggested that infants with later Autism might rely more on lower-level cues and therefore may not perform worse than neurotypical controls. Parsons et al. (2019) have previously discussed this possibility in terms of endogenous vs. exogenous influence. This could potentially explain seemingly typical early development in infants later diagnosed with Autism.

Taken together, in contrast to the notion that gaze following is affected in Autistic children early in life, findings are inconsistent. While there is some evidence indicating co-occurring effects of Autism on gaze following once a diagnosis is established, we still do not completely understand the development of gaze following in Autism. Although diminished gaze following might serve well as a diagnostic feature, gaze following during infancy may not have any clinical relevance in relation to a later Autism diagnosis.

8. Question V: How does gaze following relate to perspective-taking?

A commonly held notion that can be traced back to Scaife and Bruner (1975) is that joint attention in general, and gaze following in particular, can be characterized as a precursor to or a rudimentary form of perspective-taking (Baron-Cohen, 1995; Emery, 2000; Flom, Lee, & Muir, 2017; Moore & Corkum, 1994; Shepherd, 2010; Tomasello, 1995; Zuberbühler, 2008). This notion is implicit in the term as gaze corresponds to the subjective direction of visual attention. Thus, following gaze suggests an ability to interpret others' subjective experience as separate from one's own (e.g., Tomasello, 1995). The idea that infants can extrapolate where others are looking through perspective-taking further suggests that they possess the capacity to understand the relationship between internal states of others and external objects, and that they recognize that others have their own subjective experience in a similar way as the infant has its own experience. In this way, gaze following is conceptually similar to Theory of Mind (ToM). The reversed formulation highlights this link even more clearly: How could it be possible to follow someone's gaze (something that is referential and intentional) without a theory of mind? This issue is partially covered in the theoretical discussion concerning the mechanisms of emerging gaze following (Question II), though it deserves to be treated separately since its scope goes beyond the theoretical question of emergence; for example, if emerging gaze following relies on mechanisms other than perspective taking, it could still be a candidate precursor to perspective-taking (Charman et al., 2000), a notion compatible with all broad theoretical perspectives. Both of these views predict a correlation between gaze following and perspective taking.

Surprisingly, given the strong theoretical association, only a few empirical attempts to link infants' gaze following to perspective-taking appear in the literature. In one study, Brooks and Meltzoff (2015) found a link between gaze following at 10.5 months and later use of mental state words at 2.5 years, which in turn predicted ToM at 4.5 years. However, gaze following and ToM did not correlate. Thus, gaze following and ToM were only related through independent correlations with mental state language development. In another study, also investigating the relation between gaze following (9 and 12 months) and mental state language (desire words at 24 months and cognition words at 36 months), Kristen, Sodian, Thoermer, and Perst (2011) found no association between these abilities at any time point. However, other joint attention measures did correlate with the later use of mental state words: Point following at 9 months (not 7) predicted the use of both desire words and cognition words. At 12 months (not 15), declarative pointing predicted the use of desire words, and finally, imperative point comprehension at 12 months and imperative pointing at 12 months (not 15 or 18) also predicted desire words.

Investigating a wide range of early emerging social abilities, Charman et al. (2000) reported that the number of gaze shifts/alterations between experimenter and toy, at 20 months of age, predicted later ToM, at 44 months of age. This effect was not found in zero-order correlation but partial correlation controlling for IQ. Note that the joint attention metric they used captured IJA rather than RJA, according to the authors' description of the task. In another study, also using IJA (pointing) rather than RJA, Sodian and Kristen-Antonow (2015) found a link between joint attention at 12 months and ToM (false belief) at 50 months. Finally, investigating the impact of sex chromosome trisomies, Bouw, Swaab, Tartaglia, and van Rijn (2021) found a correlation between RJA (gaze, or point, or gaze + point following) and ToM across development from 1 to 7 years. Unfortunately, they only report the correlation in the sex chromosome trisomies group, not the control group.

Looking beyond prospective longitudinal studies, indirect evidence suggests that gaze following involves perspective-taking, at least in older infants. For example, 14- and 18-month olds follow gaze behind a transparent barrier or no barrier more frequently than an opaque barrier.

This suggests that they understand under which circumstances the interaction partner can see through the barrier, implying the presence of perspective-taking (Dunphy-Lelii & Wellman, 2004). In a study with 12-month-olds infants, Meltzoff and Brooks (2008) demonstrated that infants will use their own experience wearing a blindfold to judge whether or not to follow the gaze of people wearing a blindfold. In contrast to findings from correlational work, these two studies using more complex gaze following tasks and seem to demonstrate an implicit ability to process others' perspectives. Though research investigating a potential mentalizing capacity, or implicit ToM in infants exists in parallel to research on gaze following (see Butterfill & Apperly, 2013 for a conceptual discussion), research investigating the association between implicit ToM and gaze following is still missing.

As for now, there is not enough empirical support for the idea that early gaze following is a building block for, or early form of perspective-taking, despite this being a common notion. Longitudinal studies report no association between gaze following and later ToM. Mental state language, assumed to be important to explicit ToM, is correlated with gaze following in some studies but not in others. In contrast, older infants' capacity to complete more advanced gaze following tasks seems to imply a perspective-taking capacity. Looking at joint attention broadly, individual studies report links between various joint attention measures and ToM, but these findings need to be replicated. As it stands today, there are some indications that IJA, rather than gaze following, might be associated with ToM. However, we need more studies investigating this question if our goal is to understand the nature of emerging gaze following and its tentative link to perspective-taking. The answer to this question will have fundamental consequences for how we conceptualize gaze following, for its status in development, and for theories of emergence.

9. General discussion and summary

In this chapter, we have suggested five questions that we think are critical to address in order to advance research on gaze following. These questions represent areas of major disagreement or great uncertainty. The first question asks how developing gaze following is impacted by the society and culture in which an infant is raised. To date, there is a lack of studies targeting this question. Recent studies have demonstrated a significant impact of infants' social and emotional environment on gaze following (Astor et al., 2020). As these factors vary considerably across different cultural contexts (Bornstein, 2013; Kağitçibaşi, 2007; Keller, 2007), investigating gaze following in a variety of cultures, not only in western urban contexts, is a critical step to developing a complete understanding of this ability. Currently, there are no gaze following studies comparing infants from different cultures.

What are the mechanisms behind emerging gaze following? The second question concerns the nature of gaze following. Understanding the processes behind gaze following in infancy is essential if we want to understand its developmental path, its significance, and how it might be affected in a growing child. Despite active theoretical discussion, most studies addressing this issue are confirmatory, and truly critical research designs are rare. This is probably the most controversial question in the literature today, but perhaps also the most fundamental to answer.

Does gaze following facilitate language development? This third question targets and challenges the current narrative in the gaze following literature. This notion is constantly reiterated, and pointing to associations with later critical abilities is clearly an efficient way to highlight why research on gaze following is important. However, when looking closely at the research supporting this view, the narrative changes abruptly. Two issues are particularly noteworthy. First, in the shadow of significant correlations hides a massive number of non-significant correlations. Second, while there is a large number of studies dedicated to this question, a substantial portion confound gaze following with other RJA metrics, most commonly point following. This question is in desperate need of further investigation given how entrenched this notion currently is.

Are diminished gaze following tendencies an early marker of Autism? This fourth question has been getting some attention in the Autism literature lately (e.g., Nyström et al., 2019), though in gaze following research on typical development, the assumed link between emerging gaze following and later Autism is still frequently used as a way to highlight the clinical significance of gaze following. When reviewing the literature, it soon becomes evident that prospective infant studies fail to connect gaze following tendencies to later clinical outcome. In contrast, research on older children report differences. Thus, both in relation to language development (see above) and Autism, the field has failed to recognize that these associations, constituting cornerstones in gaze following research, rest on a theoretical rather than an empirical foundation.

How does gaze following relate to perspective-taking? The fifth and last question targets another, often implicit notion, that gaze following is related to perspective-taking, either as a manifestation of (Tomasello, 1995) or a precursor to (e.g., Charman et al., 2000) this more advanced social ability. Surprisingly few studies have addressed this question directly, and the limited data available do not indicate a clear connection between gaze following and perspective-taking. Despite this lack of association, gaze following is often used to imply perspective-taking in both infants and animals (Baron-Cohen, 1995; Emery, 2000; Shepherd, 2010; Tomasello, 1995). This question clearly needs more attention.

The five questions covered here demonstrate that the developmental path of gaze following and its role in relation to other abilities and conditions are still far from understood. It is impossible to know why the field has failed to recognize these uncertainties, but it is equally remarkable in contrast to how these questions have been handled in the literature. Could the problem be that the field as a whole is so focused on finding connections to later development that critical questions are not asked? Another suggestion is that these questions constitute a threat to the core of the gaze following research program, in which many of us are heavily invested. We carefully note that, on the one hand, there is an incentive not to pay attention to these problems. The current narrative perfectly serves as legitimizing existing research programs. On the other hand, resolving these questions has the potential to create a paradigm shift within the field. From our perspective, it is only by addressing these issues that the field will be able to move forward.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596.

Astor, K., & Gredebäck, G. (2019). Gaze following in 4.5- and 6-month-old infants: The impact of proximity on standard gaze following performance tests. *Infancy*, *24*(1), 79–89. https://doi.org/10.1111/infa.12261.

Astor, K., Lindskog, M., Forssman, L., Kenward, B., Fransson, M., Skalkidou, A., et al. (2020). Social and emotional contexts predict the development of gaze following in early infancy. *Royal Society Open Science*, 7(9), 201178. https://doi.org/10.1098/rsos.201178.

Astor, K., Thiele, M., & Gredebäck, G. (2021). Gaze following emergence relies on both perceptual cues and social awareness. *Cognitive Development*, *60*. https://doi.org/10.1016/j.cogdev.2021.101121.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. MIT Press.

Baron-Cohen, S., Baldwin, D. A., & Crowson, M. (1997). Do children with autism use the speaker's direction of gaze strategy to crack the code of language? *Child Development*, *68*(1), 48–57.

Barry-Anwar, R. A., Burris, J. L., Graf Estes, K., & Rivera, S. M. (2017). Caregivers and strangers: The influence of familiarity on gaze following and learning. *Infant Behavior & Development*, *46*, 46–58. https://doi.org/10.1016/j.infbeh.2016.11.005.

Batki, A., Baron-Cohen, S., Wheelwright, S., Connellan, J., & Ahluwalia, J. (2000). Is there an innate gaze module? Evidence from human neonates. *Infant Behavior and Development*, *23*(2), 223–229.

Bedford, R., Elsabbagh, M., Gliga, T., Pickles, A., Senju, A., Charman, T., et al. (2012). Precursors to social and communication difficulties in infants at-risk for autism: Gaze following and attentional engagement. *Journal of Autism and Developmental Disorders*, *42*(10), 2208–2218. https://doi.org/10.1007/s10803-012-1450-y.

Beuker, K. T., Rommelse, N. N. J., Donders, R., & Buitelaar, J. K. (2013). Development of early communication skills in the first two years of life. *Infant Behavior & Development*, *36*(1), 71–83. https://doi.org/10.1016/j.infbeh.2012.11.001.

Bornstein, M. H. (2013). Parenting and child mental health: A cross-cultural perspective. *World Psychiatry*, *12*(3), 258–265. https://doi.org/10.1002/wps.20071.

Bouw, N., Swaab, H., Tartaglia, N., & van Rijn, S. (2021). The impact of sex chromosome Trisomies (XXX, XXY, XYY) on early social cognition: Social orienting, joint attention, and theory of mind. *Archives of Clinical Neuropsychology*, *37*(1), 63–77. https://doi.org/10.1093/arclin/acab042.

Bräuer, J., Call, J., & Tomasello, M. (2005). All great ape species follow gaze to distant locations and around barriers. *Journal of Comparative Psychology*, *119*(2), 145–154. https://doi.org/10.1037/0735-7036.119.2.145.

Brooks, R., & Meltzoff, A. N. (2002). The importance of eyes: How infants interpret adult looking behavior. *Developmental Psychology*, *38*(6), 958–966. https://doi.org/10.1037/0012-1649.38.6.958.

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science*, *8*(6), 535–543. https://doi.org/10.1111/j.1467-7687.2005.00445.x.

Brooks, R., & Meltzoff, A. N. (2008). Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: A longitudinal, growth curve modeling study. *Journal of Child Language*, *35*(1), 207–220. https://doi.org/10.1017/S030500090700829X.

Brooks, R., & Meltzoff, A. N. (2015). Connecting the dots from infancy to childhood: A longitudinal study connecting gaze following, language, and explicit theory of mind. *Journal of Experimental Child Psychology*, *130*, 67–78. https://doi.org/10.1016/j.jecp.2014.09.010.

Brooks, R., Singleton, J. L., & Meltzoff, A. N. (2020). Enhanced gaze-following behavior in deaf infants of deaf parents. *Developmental Science*, *23*(2), e12900-n/a. https://doi.org/10.1111/desc.12900.

Butterfill, S. A., & Apperly, I. A. (2013). How to construct a minimal theory of mind. *Mind & Language*, *28*(5), 606–637.

Butterworth, G. (1991). The ontogeny and phylogeny of joint visual attention. In A. Whiten (Ed.), *Natural theories of mind: Evolution, development and simulation of everyday mindreading* (pp. 223–232). Basil Blackwell.

Butterworth, G., & Cochran, E. (1980). Towards a mechanism of joint visual attention in human infancy. *International Journal of Behavioral Development*, *3*, 253–272. https://doi.org/10.1177/016502548000300303.

Butterworth, G., & Jarrett, N. (1991). What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology*, *9*(1), 55–72. https://doi.org/10.1111/j.2044-835X.1991.tb00862.x.

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, *i-174*.

Carpenter, M., Pennington, B. F., & Rogers, S. J. (2002). Interrelations among social-cognitive skills in young children with autism. *Journal of Autism and Developmental Disorders*, *32*(2), 91–106.

Çetinçelik, B. M., Rowland, C. F., & Snijders, T. M. (2021). Do the eyes have it? A systematic review on the role of eye gaze in infant language development. *Frontiers in Psychology*, *11*, 589096. https://doi.org/10.3389/fpsyg.2020.589096.

Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Cox, A., & Drew, A. (2000). Testing joint attention, imitation, and play as infancy precursors to language and theory of mind. *Cognitive Development*, *15*(4), 481–498.

Chow, V., Poulin-Dubois, D., & Lewis, J. (2008). To see or not to see: Infants prefer to follow the gaze of a reliable looker. *Developmental Science*, *11*(5), 761–770. https://doi.org/10.1111/j.1467-7687.2008.00726.x.

Christensen, D. L., Bilder, D. A., Zahorodny, W., Pettygrove, S., Durkin, M. S., Fitzgerald, R. T., et al. (2016). Prevalence and characteristics of autism spectrum disorder among 4-year-old children in the autism and developmental disabilities monitoring network. *Journal of Developmental and Behavioral Pediatrics*, *37*(1), 1–8. https://doi.org/10.1097/DBP.0000000000000235.

Cohen, A. S., Sasaki, J. Y., German, T. C., & Kim, H. S. (2017). Automatic mechanisms for social attention are culturally penetrable. *Cognitive Science*, *41*(1), 242–258. https://doi.org/10.1111/cogs.12329.

Corkum, V., & Moore, C. (1998). Origins of joint visual attention in infants. *Developmental Psychology*, *34*(1), 28–38. https://doi.org/10.1037//0012-1649.34.1.28.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 148–153. https://doi.org/10.1016/j.tics.2009.01.005.

De Schuymer, L., De Groote, I., Striano, T., Stahl, D., & Roeyers, H. (2011). Dyadic and triadic skills in preterm and full term infants: A longitudinal study in the first year. *Infant Behavior & Development*, *34*(1), 179–188. https://doi.org/10.1016/j.infbeh.2010.12.007.

Deák, G. O. (2015). When and where do infants follow gaze? In *Paper presented at the IEEE* (pp. 182–187). https://doi.org/10.1109/DEVLRN.2015.7346138.

Deák, G. O., Flom, R. A., & Pick, A. D. (2000). Effects of gesture and target on 12- and 18-month-olds' joint visual attention to objects in front of or behind them. *Developmental Psychology*, *36*(4), 511–523. https://doi.org/10.1037/0012-1649.36.4.511.

Deák, G. O., Krasno, A. M., Triesch, J., Lewis, J., & Sepeta, L. (2014). Watch the hands: Infants can learn to follow gaze by seeing adults manipulate objects. *Developmental Science*, *17*(2), 270–281. https://doi.org/10.1111/desc.12122.

Dehaene, S. (2020). *How we learn: The New Science of Education and the Brain*. UK: Penguin.

Del Bianco, T., Falck-Ytter, T., Thorup, E., & Gredebäck, G. (2019). The developmental origins of gaze-following in human infants. *Infancy*. https://doi.org/10.1111/infa.12276.

Deligianni, F., Senju, A., Gergely, G., & Csibra, G. (2011). Automated gaze-contingent objects elicit orientation following in 8-month-old infants. *Developmental Psychology*, *47*(6), 1499.

D'Entremont, B. (2000). A perceptual–attentional explanation of gaze following in 3- and 6- month-olds. *Developmental Science*, *3*(3), 302–311. https://doi.org/10.1111/1467-7687.00124.

D'Entremont, B., Hains, S. M. J., & Muir, D. W. (1997). A demonstration of gaze following in 3- to 6-month-olds. *Infant Behavior & Development*, *20*(4), 569–572. https://doi.org/10.1016/S0163-6383(97)90048-5.

Dunphy-Lelii, S., & Wellman, H. M. (2004). Infants' understanding of occlusion of others' line-of-sight: Implications for an emerging theory of mind. *European Journal of Developmental Psychology*, *1*(1), 49–66.

Elsabbagh, M., Fernandes, J., Jane Webb, S., Dawson, G., Charman, T., Johnson, M. H., et al. (2013). Disengagement of visual attention in infancy is associated with emerging autism in toddlerhood. *Biological Psychiatry*, *74*(3), 189. https://doi.org/10.1016/j.biopsych.2012.11.030.

Emery, N. J. (2000). The eyes have it: The neuroethology, function and evolution of social gaze. *Neuroscience & Biobehavioral Reviews*, *24*(6), 581–604.
Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(14), 9602–9605. https://doi.org/10.1073/pnas.152159999.
Farroni, T., Johnson, M. H., Brockbank, M., & Simion, F. (2000). Infants' use of gaze direction to cue attention: The importance of perceived motion. *Visual Cognition*, *7*(6), 705–718. https://doi.org/10.1080/135062800501443999.
Farroni, T., Mansfield, E. M., Lai, C., & Johnson, M. H. (2003). Infants perceiving and acting on the eyes: Tests of an evolutionary hypothesis. *Journal of Experimental Child Psychology*, *85*(3), 199–212. https://doi.org/10.1016/s0022-0965(03)00022-5.
Farroni, T., Massaccesi, S., Pividori, D., & Johnson, M. H. (2004). Gaze following in newborns. *Infancy*, *5*(1), 39–60. https://doi.org/10.1207/s15327078in0501_2.
Fasel, I., Deák, G. O., Triesch, J., & Movellan, J. (2002). Combining embodied models and empirical research for understanding the development of shared attention. In *Proceedings of the 2nd International Conference on Development and Learning. ICDL 2002* (pp. 21–27). IEEE.
Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., et al. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*. i-185.
Filipek, P. A., Accardo, P. J., Baranek, G. T., Cook, E. H., Dawson, G., Gordon, B., et al. (1999). The screening and diagnosis of autistic spectrum disorders. *Journal of Autism and Developmental Disorders*, *29*(6), 439–484.
Flom, R., Lee, K., & Muir, D. (2017). *Gaze-following: Its development and significance*. Psychology Press.
Gillespie-Lynch, K., Elias, R., Escudero, P., Hutman, T., & Johnson, S. P. (2013). Atypical gaze following in autism: A comparison of three potential mechanisms. *Journal of Autism and Developmental Disorders*, *43*(12), 2779–2792. https://doi.org/10.1007/s10803-013-1818-7.
Gliga, T., Elsabbagh, M., Hudry, K., Charman, T., Johnson, M. H., & BASIS team. (2012). Gaze following, gaze reading, and word learning in children at risk for autism. *Child Development*, *83*(3), 926–938.
Gredebäck, G., Astor, K., & Fawcett, C. (2018). Gaze following is not dependent on ostensive cues: A critical test of natural pedagogy. *Child Development*, *89*(6), 2091–2098. https://doi.org/10.1111/cdev.13026.
Gredebäck, G., Fikke, L., & Melinder, A. (2010). The development of joint visual attention: A longitudinal study of gaze following during interactions with mothers and strangers. *Developmental Science*, *13*(6), 839–848. https://doi.org/10.1111/j.1467-7687.2009.00945.x.
Gredebäck, G., Johnson, S., & von Hofsten, C. (2010). Eye tracking in infancy research. *Developmental Neuropsychology*, *35*(1), 1–19. https://doi.org/10.1080/87565640903325758.
Hernik, M., & Broesch, T. (2019). Infant gaze following depends on communicative signals: An eye-tracking study of 5- to 7-month-olds in vanuatu. *Developmental Science*, *22*(4), e12779-n/a. https://doi.org/10.1111/desc.12779.
Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. J., Brown, E., Chung, H. L., et al. (2000). Breaking the language barrier: An emergentist coalition model of the origins of word learning. *Monographs of the Society for Research in Child Development*, *65*(3). x1–122.
Ishikawa, M., Yoshimura, M., Sato, H., & Itakura, S. (2019). Effects of attentional behaviours on infant visual preferences and object choice. *Cognitive Processing*, *20*(3), 317–324. https://doi.org/10.1007/s10339-019-00918-x.
Itakura, S. (1996). An exploratory study of gaze-monitoring in non-human primates 1. *Japanese Psychological Research*, *38*(3), 174–180.

Jasso, H., Triesch, J., Deák, G., & Lewis, J. M. (2012). A unified account of gaze following. *IEEE Transactions on Autonomous Mental Development*, *4*(4), 257–272.
Juvrud, J., Bakker, M., Kaduk, K., DeValk, J. M., Gredebäck, G., & Kenward, B. (2019). Longitudinal continuity in understanding and production of giving-related behavior from infancy to childhood. *Child Development*, *90*(2), e182–e191.
Kağitçibaşi, Ç. (2007). *Family, self, and human development across cultures: Theories and applications* (2nd ed.). Lawrence Erlbaum Associates Publishers.
Kaminski, J., Riedel, J., Call, J., & Tomasello, M. (2005). Domestic goats, Capra hircus, follow gaze direction and use social cues in an object choice task. *Animal Behaviour*, *69*(1), 11–18.
Keller, H. (2007). *Cultures of infancy*. Lawrence Erlbaum Associates Publishers.
Keller, H. (2018). Universality claim of attachment theory: Children's socioemotional development across cultures. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(45), 11414–11419. https://doi.org/10.1073/pnas.1720325115.
Kesselring, T., & Müller, U. (2011). The concept of egocentrism in the context of Piaget's theory. *New Ideas in Psychology*, *29*(3), 327–345. https://doi.org/10.1016/j.newideapsych.2010.03.008.
Kotchabhakdi, N. J., Winichagoon, P., Smitasiri, S., Dhanamitta, S., & Valyasevi, A. (1987). The integration of psychosocial components in nutrition education in northeastern Thai villages. *Asia-Pacific Journal of Public Health*, *1*(2), 16–25. https://doi.org/10.1177/101053958700100205.
Kristen, S., Sodian, B., Thoermer, C., & Perst, H. (2011). Infants' joint attention skills predict toddlers' emerging mental state language. *Developmental Psychology*, *47*(5), 1207–1219. https://doi.org/10.1037/a0024808.
Landa, R. J., Holman, K. C., & Garrett-Mayer, E. (2007). Social and communication development in toddlers with early and later diagnosis of autism spectrum disorders. *Archives of General Psychiatry*, *64*(7), 853–864. https://doi.org/10.1001/archpsyc.64.7.853.
Leekam, S., Baron-Cohen, S., Perrett, D., Milders, M., & Brown, S. (1997). Eye-direction detection: A dissociation between geometric and joint attention skills in autism. *British Journal of Developmental Psychology*, *15*(1), 77–95.
Leekam, S. R., Hunnisett, E., & Moore, C. (1998). Targets and cues: Gaze-following in children with autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *39*(7), 951–962.
Markus, J., Mundy, P., Morales, M., Delgado, C. E. F., & Yale, M. (2000). Individual differences in infant skills as predictors of child-caregiver joint attention and language. *Social Development (Oxford, England)*, *9*(3), 302–315. https://doi.org/10.1111/1467-9507.00127.
Meltzoff, A. N. (2007). "like me": A foundation for social cognition. *Developmental Science*, *10*(1), 126.
Meltzoff, A. M., & Brooks, R. (2007). Eyes wide shut: The importance of eyes in infant gaze-following and understanding other minds. In R. Flom, K. Lee, & D. Muir (Eds.), *Gaze-following: Its development and significance* (pp. 217–241). Lawrence Erlbaum Associates Publishers.
Meltzoff, A. N., & Brooks, R. (2008). Self-experience as a mechanism for learning about others: A training study in social cognition. *Developmental Psychology*, *44*(5), 1257–1265. https://doi.org/10.1037/a0012888.
Michel, C., Kayhan, E., Pauen, S., & Hoehl, S. (2021). Effects of reinforcement learning on gaze following of gaze and head direction in early infancy: An interactive eye-tracking study. *Child Development*, *92*(4), e364–e382. https://doi.org/10.1111/cdev.13497.
Moore, C. (2008). The development of gaze following. *Child Development Perspectives*, *2*(2), 66–70. https://doi.org/10.1111/j.1750-8606.2008.00052.x.

Moore, C., Angelopoulos, M., & Bennett, P. (1997). The role of movement in the development of joint visual attention. *Infant Behavior and Development*, *20*(1), 83–92. https://doi.org/10.1016/S0163-6383(97)90063-1.
Moore, C., & Corkum, V. (1994). Social understanding at the end of the first year of life. *Developmental Review*, *14*(4), 349–372. https://doi.org/10.1006/drev.1994.1014.
Morales, M., Mundy, P., Delgado, C. E. F., Yale, M., Messinger, D., Neal, R., et al. (2000a). Responding to joint attention across the 6- through 24-month age period and early language acquisition. *Journal of Applied Developmental Psychology*, *21*(3), 283–298. https://doi.org/10.1016/S0193-3973(99)00040-4.
Morales, M., Mundy, P., Delgado, C. E. F., Yale, M., Neal, R., & Schwartz, H. K. (2000b). Gaze following, temperament, and language development in 6-month-olds: A replication and extension. *Infant Behavior & Development*, *23*(2), 231–236. https://doi.org/10.1016/S0163-6383(01)00038-8.
Morales, M., Mundy, P., & Rojas, J. (1998). Following the direction of gaze and language development in 6-month-olds. *Infant Behavior & Development*, *21*(2), 373–377. https://doi.org/10.1016/S0163-6383(98)90014-5.
Morissette, P., Ricard, M., & Gouin Décarie, T. G. (1995). Joint visual attention and pointing in infancy: A longitudinal study of comprehension. *British Journal of Developmental Psychology*, *13*, 163–175.
Mundy, P. (2016). *Autism and joint attention: Development, neuroscience, and clinical fundamentals*. New York: The Guilford Publications.
Mundy, P., Block, J., Delgado, C., Pomares, Y., Van Hecke, A. V., & Parlade, M. V. (2007). Individual differences and the development of joint attention in infancy. *Child Development*, *78*(3), 938–954. https://doi.org/10.1111/j.1467-8624.2007.01042.x.
Mundy, P., Fox, N., & Card, J. (2003). EEG coherence, joint attention and language development in the second year. *Developmental Science*, *6*(1), 48–54. https://doi.org/10.1111/1467-7687.00253.
Mundy, P., & Gomes, A. (1998). Individual differences in joint attention skill development in the second year. *Infant Behavior & Development*, *21*(3), 469–482. https://doi.org/10.1016/S0163-6383(98)90020-0.
Nsamenang, A. B., & Lamb, M. E. (2014). Socialization of Nso children in the Bamenda grassfields of Northwest Cameroon. In *Cross-cultural roots of minority child development* (pp. 135–149). Psychology Press.
Nygren, G., Cederlund, M., Sandberg, E., Gillstedt, F., Arvidsson, T., Carina Gillberg, I., et al. (2012). The prevalence of autism spectrum disorders in toddlers: A population study of 2-year-old Swedish children. *Journal of Autism and Developmental Disorders*, *42*(7), 1491–1497. https://doi.org/10.1007/s10803-011-1391-x.
Nyström, P., Thorup, E., Bölte, S., & Falck-Ytter, T. (2019). Joint attention in infancy and the emergence of autism. *Biological Psychiatry (1969)*, *86*(8), 631–638. https://doi.org/10.1016/j.biopsych.2019.05.006.
Okumura, Y., Kanakogi, Y., Kobayashi, T., & Itakura, S. (2017). Individual differences in object-processing explain the relationship between early gaze-following and later language development. *Cognition*, *166*, 418–424.
Ozonoff, S., Iosif, A., Baguio, F., Cook, I. C., Hill, M., Hutman, T., et al. (2010). A prospective study of the emergence of early behavioral signs of autism. *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*(3), 256–266e2. https://doi.org/10.1097/00004583-201003000-00009.
Ozonoff, S., Young, G. S., Landa, R. J., Brian, J., Bryson, S., Charman, T., et al. (2015). Diagnostic stability in young children at risk for autism spectrum disorder: A baby siblings research consortium study. *Journal of Child Psychology and Psychiatry*, *56*(9), 988–998. https://doi.org/10.1111/jcpp.12421.

Parsons, J. P., Bedford, R., Jones, E. J., Charman, T., Johnson, M. H., Gliga, T., et al. (2019). Gaze following and attention to objects in infants at familial risk for ASD. *Frontiers in Psychology*, *10*(July), 1799. https://doi.org/10.3389/fpsyg.2019.01799.

Peña, M., Arias, D., & Dehaene-Lambertz, G. (2014). Gaze following is accelerated in healthy preterm infants. *Psychological Science*, *25*(10), 1884–1892. https://doi.org/10.1177/0956797614544307.

Piaget, J. (1952). *Origins of intelligence in children*. WW Norton & Co.

Rozga, A., Hutman, T., Young, G. S., Rogers, S. J., Ozonoff, S., Dapretto, M., et al. (2011). Behavioral profiles of affected and unaffected siblings of children with autism: Contribution of measures of Mother–Infant interaction and nonverbal communication. *Journal of Autism and Developmental Disorders*, *41*(3), 287–301. https://doi.org/10.1007/s10803-010-1051-6.

Sacrey, L. R., Bryson, S. E., & Zwaigenbaum, L. (2013). Prospective examination of visual attention during play in infants at high-risk for autism spectrum disorder: A longitudinal study from 6 to 36 months of age. *Behavioural Brain Research*, *256*, 441–450. https://doi.org/10.1016/j.bbr.2013.08.028.

Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature*, *253*(5489), 265–266. https://doi.org/10.1038/253265a0.

Schloegl, C., Kotrschal, K., & Bugnyar, T. (2007). Gaze following in common ravens, Corvus corax: Ontogeny and habituation. *Animal Behaviour*, *74*(4), 769–778.

Senju, A., Vernetti, A., Ganea, N., Hudry, K., Tucker, L., Charman, T., et al. (2015). Early social experience affects the development of eye gaze processing. *Current Biology*, *25*(23), 3086–3091.

Shepherd, S. V. (2010). Following gaze: Gaze-following behavior as a window into social cognition. *Frontiers in Integrative Neuroscience*, *4*, 5.

Silverstein, P., Feng, J., Westermann, G., Parise, E., & Twomey, K. E. (2021). Infants learn to follow gaze in stages: Evidence confirming a robotic prediction. *Open Mind: Discoveries in Cognitive Science*, *5*, 174–188. https://doi.org/10.1162/opmi_a_00049.

Slaughter, V., & McConnell, D. (2003). Emergence of joint attention: Relationships between gaze following, social referencing, imitation, and naming in infancy. *The Journal of Genetic Psychology*, *164*(1), 54–71. https://doi.org/10.1080/00221320309597503.

Sodian, B., & Kristen-Antonow, S. (2015). Declarative joint attention as a foundation of theory of mind. *Developmental Psychology*, *51*(9), 1190–1200. https://doi.org/10.1037/dev0000039.

Szufnarowska, J., Rohlfing, K. J., Fawcett, C., & Gredebäck, G. (2014). Is ostension any more than attention? *Scientific Reports*, *4*, 5304. https://doi.org/10.1038/srep05304.

Téglás, E., Gergely, A., Kupán, K., Miklósi, Á., & Topál, J. (2012). Dogs' gaze following is tuned to human communicative signals. *Current Biology*, *22*(3), 209–212.

Tenenbaum, E. J., Sobel, D. M., Sheinkopf, S. J., Malle, B. F., & Morgan, J. L. (2015). Attention to the mouth and gaze following in infancy predict language development. *Journal of Child Language*, *42*(6), 1173–1190. https://doi.org/10.1017/S0305000914000725.

Thorup, E., Nyström, P., Bölte, S., & Falck-Ytter, T. (2021). What are you looking at? Gaze following with and without target objects in ASD and typical development. Autism. *The International Journal of Research and Practice*. https://doi.org/10.1177/13623613211061940. 136236132110619-13623613211061940.

Thorup, E., Nyström, P., Gredebäck, G., Bölte, S., & Falck-Ytter, T. (2016). *Altered gaze following during live interaction in infants at risk for autism: An eye tracking study*. https://doi.org/10.1186/s13229-016-0069-9.

Tomasello, M. (1995). Joint attention as social cognition. *Joint Attention: Its Origins and Role in Development*, *103130*, 103–130.

Tomasello, M., Call, J., & Hare, B. (1998). Five primate species follow the visual gaze of conspecifics. *Animal Behaviour*, *55*(4), 1063–1069. https://doi.org/10.1006/anbe.1997.0636.

Tomasello, M., Hare, B., Lehmann, H., & Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: The cooperative eye hypothesis. *Journal of Human Evolution*, *52*(3), 314–320. https://doi.org/10.1016/j.jhevo1.2006.10.001.

Triesch, J., Teuscher, C., Deák, G. O., & Carlson, E. (2006). Gaze following: Why (not) learn it? *Developmental Science*, *9*(2), 125–147. https://doi.org/10.1111/j.1467-7687.2006.00470.x.

von Hofsten, C., Dahlström, E., & Fredriksson, Y. (2005). 12-month-old infants' perception of attention direction in static video images. *Infancy*, *8*(3), 217–231. https://doi.org/10.1207/s15327078in0803_2.

Zuberbühler, K. (2008). Audience effects. *Current Biology*, *18*(5), R189–R190.

This page intentionally left blank

CHAPTER EIGHT

Young children's cooperation and conflict with other children

Dale F. Hay*, Amy Paine, and Charlotte Robinson
Cardiff University, Cardiff, United Kingdom
*Corresponding author: e-mail address: haydf@cardiff.ac.uk

Contents

Abstract

Early forms of cooperation and conflict feature regularly in young children's interactions with other people. However, these two types of social interaction are only rarely studied together in the same sample. In this chapter we review studies of cooperation and conflict in children under 3 years of age, with a particular focus on peer interaction. Only a few studies examined cooperation and conflict in parallel. To illustrate how conflict and cooperation can be studied simultaneously, we present findings from a longitudinal study of social development, in which previously unacquainted toddlers were observed during laboratory birthday parties. These analyses revealed that the two types of interaction are positively associated and provide opportunities for young children to refine their social skills.

Our chapter reports findings from studies of very young children's cooperation and conflict with other people, with a particular focus on interactions between same-aged peers. By focusing on peer interaction, we are able to examine early forms of cooperation and conflict that occur between young equals, as opposed to studying compliance and disciplinary encounters in

Advances in Child Development and Behavior, Volume 63
ISSN 0065-2407
https://doi.org/10.1016/bs.acdb.2022.04.004

hierarchical relationships with parents (Hinde, 1979). This allows us to examine the children's developing social competence, not scaffolded by adults' behavior.

We shall examine the evidence for both cooperation and conflict in the social interactions of children under 3 years of age. After reviewing relevant studies, we shall provide an example of how cooperation and conflict between peers can be observed in the same study by drawing upon previously unpublished findings from our own longitudinal study of early social development (Hay et al., 2021). First, however, it is important to place these studies within the broader literature on cooperative, competition, and conflict.

Within social psychology, cooperation was often studied in conjunction with competition, in observational and experimental studies of social groups. For example, in social interdependence theory (Deutsch, 1949; Lewin, 1948), cooperation and competition are considered to be equally important forms of interaction within social groups; in this perspective, both cooperation and competition are seen as social acts that can be distinguished from individualistic behavior (Johnson & Johnson, 2011).

Children's cooperation and competition were sometimes studied together in experimental paradigms where the children could choose to cooperate or compete (e.g., Charlesworth, 1996; Green & Rechis, 2006; Madsen, 1967). More recently, however, children's cooperation is less often studied in parallel with competition or conflict. Perhaps this is because developmental scientists often focus on long term change and continuity in individuals' temperaments, skills and behaviors and pay less attention to the dynamics of social interactions in the short term. Thus, in developmental psychology, when children's conflicts are studied, they are often seen as a framework in which to study individual children's aggression (e.g., Rubin, Burgess, Dwyer, & Hastings, 2003).

Similarly, when cooperation is studied at the individual level, it may be seen as one of several categories of prosocial behavior (e.g., Radke-Yarrow, Zahn-Waxler, & Chapman, 1983), although some developmental theorists do not consider cooperation to meet the theoretical criteria for prosocial action (e.g., Eisenberg, Fabes, & Spinrad, 2006). In the latter perspective, where prosocial behavior implies active concern for another person rather than the self, cooperation does not qualify as prosocial, because the child also benefits. Perhaps because of this difference in definitional criteria, cooperation is not always included in operational definitions of prosocial behavior.

At the individual level of analysis, aggression and prosocial behavior are themselves only rarely studied together in the same sample (for a review, see Hay et al., 2021). Similarly, in the literature on early social development, cooperation and conflict are not often examined in relation to each other, and so distinct literatures have emerged. In this chapter, our aim is to integrate the different sets of findings on cooperation and conflict, in line with the tradition of social interdependence theory, which sees them as interlinked (Deutsch, 1949). If so, they might be seen as opposite ways of dealing with the social world or, alternatively, as two forms of social engagement that happen to be positively associated. First, however, we begin by focusing on each topic in turn.

1. Cooperation in the first 3 years

In the Oxford English Dictionary, *cooperate* is defined simply as "to work together." The dictionary definition implies that some effort must be put in by two or more parties to sustain cooperation over time. In the case of children under 3 years of age, "work" can be defined as the efforts they put into sustaining a harmonious interaction over a few minutes. In other words, young children's cooperative play takes quite a bit of work.

The Oxford dictionary definition's emphasis on working *together* helps to distinguish between cooperation and compliance. Simply obeying orders or giving in to another's demands does not imply that both parties are working together toward a mutual goal. Sustaining a pleasant interaction can be a goal in its own right, but in many cases cooperating partners are trying to solve a problem or attain an external goal. Therefore, in this chapter, we examine the evidence for both types of cooperation seen in the first years of life: *cooperative play* and *cooperative problem solving*.

1.1 Cooperative play

Social games with parents. The development of a capacity for cooperative play begins in the early months of life, when infants interact with their caregivers. Early social games such as 'peek-a-boo' have long been studied in the context of infants' interactions with their caregivers in the first year (e.g., Garvey, 1977; Trevarthen & Reddy, 2017), which can be detected in the first months after birth (e.g., Fantasia, Fasulo, Costall, & López, 2014). Over the first year, infants gradually take on a more active role in such games, moving from early games where the infant's only role is to smile or laugh to games in which the infants hide their own faces from their companions (Bruner & Sherwood, 1976). Infants' active contributions to games with their parents, which includes teasing and clowning around (Reddy & Mireault, 2015), reveals their cooperative playfulness.

Cooperative play with peers. However, to examine the development of infants' and toddlers' own cooperative skills, it is particularly helpful to observe their interactions with peers. Peers of similar age usually have similar levels of social competence but little power, and so the cooperative interactions they engage in are not directed by more competent or powerful adults (Verba, 1994). Thus, studies of peer interaction provide opportunities to distinguish between cooperation with other children and compliance with parents and teachers. In some studies of young peers, the children are observed in familiar settings, at home or their usual child care settings. In other research, young children are introduced to new acquaintances in laboratory playrooms.

Classic studies of young children's cooperative play with peers were conducted in nursery schools in the 1930s. Some studies suggested that cooperative play was relatively rare in the preschool age range, compared to solitary and parallel play, where children sat near each other, not working together, but rather pursuing their own interests (Parten, 1932). However, other studies have shown that cooperative play with peers can be observed even before children attend preschool.

For example, cooperative play between peers has been documented in the second year, when toddlers began to engage in mutual games with peers (Goldman & Ross, 1978; Ross & Goldman, 1977). Ross and her colleagues specified four criteria that defined mutual games: (1) *mutual engagement* in the interaction; (2) *repetition* of key actions (such as rolling a toy toward the play partner) which defined the content of the game; (3) *alternation of turns*; and *non-literality*, a sense of playful unreality that distinguished such mutual games from other, more serious types of interactions. The first three criteria measure the cooperation that is being sustained by both partners; the fourth criterion highlights the fact that social games are intrinsically playful. Subsequent play researchers specified that the repetition in such games must be playful, not a form of stereotypy (Graham, Burghardt, & Wiens, 2010).

Other investigators explored the processes that underlie cooperative play with peers. For example, Eckerman and colleagues noted that social games defined by Ross's criteria are relatively rare; they emphasized the importance of nonverbal imitation in infants' attempts to coordinate their actions with those of other infants, even if such coordination did not meet all the criteria for a mutual game (Eckerman, Davis, & Didow, 1989).

It is noteworthy that such coordinated interactions between peers transcend culture; Eckerman and her colleagues discovered that imitation helped infants coordinate their actions with others in both US and Indonesian samples (Eckerman et al., 1989; Eckerman & Whitehead, 1999). However, a longitudinal study in the US of peers' interactions from 16 to 32 months revealed that, as the toddlers grew older, pure imitation gave way to complementary actions as a means of sustaining mutual interaction with peer partners (Eckerman et al., 1989).

For the most part, studies of cooperative play between peers have focused on relatively small, less representative samples. However, in the large, nationally representative NICHD child care sample in the US, 612 36-month-olds were each observed with a familiar peer, who was of the same gender as the focal child (NICHD Early Child Care Research Network, 2001). Demographic analyses showed that, in comparison with the whole sample, those parents who were able to arrange the observational sessions were more affluent, better educated, and more likely to live in two-parent families. Nevertheless, the NICHD subsample provides one of the largest set of observations of early peer interaction with a standardized procedure in a familiar setting (the usual child care setting or the home). Positive social interactions, including cooperation, were recorded, with physical and verbal aggression recorded as well. The children's positive social skills shown during the peer interaction were associated with verbal ability and mothers' sensitivity, with girls showing more positive sociability with peers than boys did. However, the girls were also more self-assertive during the peer interaction than were boys. These data suggest that while the capacity for skillful interactions with peers is present by the third birthday, there are significant individual differences as well.

In the NICHD study, subsequent analyses of the peer observation at 36 months revealed that the children's cooperative play with the peer was positively associated with another form of prosocial behavior, the focal child's expressions of concern for the peer (Blandon & Scrimgeour, 2015). However, these two types of prosocial behavior had somewhat different correlates. A child's cooperation and concern for a peer's well-being were both associated with the mother's sensitivity. However, cooperation was correlated with the immediate situation, that is, peer's positive behavior during the observation, whereas concern for the peer was affected by the quality of the peers' previous relationship. These findings confirm that cooperative play and concern for others both qualify as prosocial behaviors, but cooperation is more influenced by the peer's immediate contributions to the present interaction. This finding underscores the fact that cooperation is by its very nature dyadic, not an individual act of concern or kindness.

1.2 Cooperative problem solving

A number of studies of young children's cooperation have not focused on peer interaction, but rather on children's responses to experimental tasks. In recent years, cooperation has been studied within both an evolutionary and a developmental framework (for a review, see Tomasello & Vaish, 2013). The studies reviewed in the previous section have documented the abilities of 2- and 3-year-old children to collaborate with peers and, moreover, to express enjoyment in their mutual games. Young children's pleasure in collaboration stands in contrast to the preferences of the great apes, who are able to collaborate if need be, but prefer to work on their own (Tomasello & Vaish, 2013).

This fusion of developmental and evolutionary perspectives on cooperation has led to comparisons of young children's abilities to cooperate with the cooperative abilities of chimpanzees (e.g., Warneken, Chen, & Tomasello, 2006). To create fair comparisons, cooperation has been measured more formally in experimental paradigms where a child (or chimpanzee) must work together with another to reach a mutual goal. Even beyond the comparison with other species, experimental measures of children's cooperative problem-solving provide opportunities to chart the development of the ability to work with another person toward a mutual goal. In much past research, young children's cooperative abilities have been assessed with adult experimenters (e.g., Kartner, Schumachker, & Collard, 2014; Meyer, Bekkering, Haartsen, Stapel, & Hunnius, 2015; Warneken et al., 2006; Warneken & Tomasello, 2007). However, to make direct comparisons with the observational literature on cooperative play, as described above, it is important to focus on the extent of cooperative problem-solving between peers. In general, these studies of cooperative problem-solving have focused on age-related cooperative competence in relatively small convenience samples. The ingenious tasks that have been devised have produced broad findings about age differences, highlighting the emergence of successful coordination around the time of the third birthday.

Age differences. In a cross-sectional experiment on young children ranging from 12 to 30 months of age, the children were paired with unfamiliar peers of the same age and gender and presented with a problem that could only be solved via cooperative action (Brownell & Carriger, 1990). Three possible cooperative problems were given to the children, counterbalanced across pairs of children. All the problems required the children to use a tool (varied across the different problems) to gain access to toy animals; the toys could only be reached if the children worked together, one operating the tool that opened a window to gain access to the toys and the other moving to the position where they could reach the animals while the window was still open. Thus, these cooperation problems required the taking of complementary roles and synchrony in time. No 12-month-old solved any of the problems, and the 18-month-olds had difficulty with the tasks. However, 24- and 30-month-olds were able to coordinate their actions with those of their partners to attain the mutual goal.

A subsequent study in this laboratory probed into a similar cooperative problem-solving task given to 19-, 23- and 27-month-olds (Brownell, Ramani, & Zerwas, 2006a, 2006b). In this task, pairs of previously unacquainted children had to pull on handles either at the same time or in sequence to activate a toy dog. Only a minority of 19-month-olds were able to coordinate with the peer; in contrast, almost all of the 27-month-olds did so successfully. Such cooperative-problem solving was correlated with the children's joint attention skills, even when accounting for chronological age.

These findings suggest that cooperative problem-solving skills may begin to emerge in the second year but strengthen thereafter. In a study of 19-month-old Dutch peers, only a quarter of the dyads tested were able to coordinate their actions on a cooperation task that entailed opening a lid so that a ball could travel down and out of a tube (Hunnius, Bekkering, & Cillesen, 2009). Rather than stimulating cooperation, the ball-and-tube task actually provoked conflict in the 19-month-old pairs, with antagonistic actions being more frequent than affiliative behavior. These findings draw attention to the fluidity of peer interactions at this age, which encompass both conflict and spontaneous cooperative play, often in the same sequence of actions and reactions.

Gender and individual differences. The studies of cooperative problem-solving tasks have sought evidence for age differences, but gender differences and individual differences may also be important. For example, in a sample of children working on a similar collaboration task (coordinating actions to retrieve stickers from a tube), successful cooperation was not apparent in children younger than 2½ years of age; highly competent performance on the task was shown only by 3½-year-olds (Ashley & Tomasello, 1998). However, in that sample, age group was confounded with different patterns of gender composition, which may have introduced additional variance in the analyses.

It is likely that individual children differ in their willingness to cooperate with other child participants, as well as in their abilities to perform competently on collaboration tasks. Some sources of variation were investigated in a study of 23 pairs of toddlers (mean age 36 months) who were observed during free play, an obligatory coordination task, and a follow-up session where they could collaborate or play by themselves (Schumacher & Kärtner, 2015). Mothers had been asked to assess the toddlers' temperament, sociability and mastery motivation and predict their children's behavior during the coordination task. During the obligatory task, children spent about half their time attempting to coordinate with the peer; in the subsequent follow-up session, they spent far less time doing so but did engage in a sociable form of parallel play, working on the tasks on their own while otherwise engaging with the other child. Those pairs who had succeeded on the coordination task originally (50% of the sample) were more likely to coordinate on the new tasks; however, those who had been rated as having high mastery motivation were less likely to do so.

What is also evident is the fact that these assessments of cooperative competence are embedded in a broader compliance task: The toddlers must work within the constraints of the experimental situation, in terms of the testing environment, the experimenter's instructions, and the particular demands of the experimental tasks. Success on the cooperative problem-solving tasks will reflect the children's testability, their receptive language (in terms of understanding the instructions), and their abilities to use and manipulate the equipment. These elements of the experimental tasks are likely to introduce measurement error into claims about age-related change in rather small samples.

When toddlers do work together, either spontaneously during free play or in response to the demands of a collaboration task, their interactions sometimes go awry. One child may resist or protest against something the peer has done; the peer may then retaliate. Peer interaction in this age range often takes the form of a set of episodes that may incorporate conflict along with cooperation. Therefore, to understand the broader context of early cooperation, it is helpful to examine what is known about very young children's conflicts with their peers.

2. Conflict in the first 3 years

2.1 Features of young children's conflicts

Conflict, which can be defined as opposition between two or more people (Shantz, 1987), is a normative feature of social life. In observational studies of children's interactions, conflict has been defined as a social interaction in which a child resists, protests, or retaliates against something another child or adult has done (Hay & Ross, 1982). Children engage in conflict with parents, siblings, and peers, beginning at an early age. However, young children's engagement in conflict should not be seen a sign of immaturity; peer conflict continues into middle childhood and adolescence (Noakes & Rinaldi, 2006).

Similarly, children's conflicts with other people should not be interpreted as a sign of maladjustment or an indication of future behavioral problems. Nor does the rate of conflict necessarily reflect the quality of the child's family relationships. For example, in a study of toddlers' conflicts with their mothers (Laible, Panfile, & Makariev, 2008), the frequency of conflict was not related to the quality of the mother–toddler relationship.

Peer conflict. Conflicts between same-aged peers, which arise on visits to other households or in child care settings, have been studied as fleeting encounters, not emotionally charged features of particular peer relationships. During the first 3 years, conflicts between peers arise most often over issues of personal space and the use of resources. Such conflicts over territory and property can be detected by the first birthday (Caplan, Vespo, Pedersen, & Hay, 1991; Hay et al., 2011) and continue to occur over the next 2 years. When toddlers spend time with other children, peer interactions sometimes feature quick shifts between cooperative play and conflict over resources.

A number of factors provoke conflict between peers. For example, in a longitudinal study of 28 Swiss children observed during free play in their day care centers at 8, 14, and 22 months (Licht, Simoni, & Perrig-Chello, 2008), the actions that triggered peer conflict were recorded over time. Over the course of the study, 98 conflicts were identified and analyzed. At 8 months, most conflict was precipitated by one infant interrupting the play of another infant; at 14 months, infants might try to take away another infant's bottle or take over other objects that the peer had been exploring. By 22 months, such conflict over resources could be interpreted as possession disputes, with the children asserting ownership of particular objects (e.g., by saying "Mine!").

Although the authors were loath to overinterpret the youngest infants' behavior as conflict over possessions, it is evident that disputes over the mutual use of space and resources are evident by the first birthday (Caplan et al., 1991). Possession issues remain a very common reason for conflict between young children; however, they also engage in conflicts about ideas and social conventions (Chen, Fein, Killen, & Tam, 2001). As children grow older, possession disputes with peers become less common; however, possession rights remain the most common reason for sibling conflict throughout middle childhood and into adolescence (Ross, Conant, & Vickar, 2011).

Conflict is not synonymous with aggression. Conflicts are co-constructed social interactions, and the participants may or may not use aggression to pursue their goals. Conflict between young peers sometimes does incorporate the use of force, but young children use force less often than other forms of protest or resistance (e.g., Ashby & Nielsen-Hewett, 2012; Caplan et al., 1991; Chen et al., 2001; Feldman, Masalha, & Derdikman-Eiron, 2010; Hay et al., 2021; Hay, Castle, & Davies, 2000).

The use of force against peers declines as children grow older; in middle childhood, children are likely to react to aggressive peers by excluding them from play or using other forms of social aggression, rather than deploying physical force (e.g., Fanger, Frankel, & Hazen, 2012). However, there are individual differences in the inclination to use force, which show stability over time. Individual differences in aggression emerge in infancy (e.g., Tremblay et al., 1999) and individual tendencies to use force during conflicts with peers are evident by 12 months of age (Hay et al., 2011). These individual differences in aggressiveness persist into middle childhood (Hay et al., 2021).

In addition to being an arena in which some children deploy force, young children's conflicts also provide a framework in which they develop other effective ways of pursuing their interests and reacting to other children's incursions on their possessions and personal space. For example, with the acquisition of language, toddlers begin to declare their possession rights by claiming 'Mine!' A short-term longitudinal study of toddlers playing with familiar peers showed that initially such claims were associated with the use of force, but those children who were able to voice their claims for objects were more likely to share resources with the same peers 6 months later (Hay, 2006). This finding suggests that, within peer relationships, young children develop ways of managing disagreements that takes into account their language skills and their growing understanding of personal space and possession rights.

Sibling conflict. Although our primary focus is on cooperation and conflict between equals, children's strategies while in conflict with peers may be influenced by their experience at home with siblings. The majority of children (e.g., an estimated 80% in Western societies) have siblings and often spend more time with them than any other family member (McHale, Updegraff, & Whiteman, 2012). Sibling relationships are "emotionally charged…defined by strong uninhibited emotions of a positive, negative, and sometimes ambivalent quality" (Howe & Recchia, 2006, p. 1). In contrast to disputes with peers, and with the exception of twin pairs, conflicts between siblings may feature an imbalance of power due to the age difference (Howe, Paine, Ross, & Recchia, 2022). They are also embedded within the broader dynamics of a family, with parents and other siblings influencing the dynamics and outcomes of sibling disputes. For example, in analyses of a large, representative birth cohort in the UK Millennium Study, in which family structure was taken into account, full siblings were significantly more likely to engage in conflict than were half-siblings (Tanskanen, Danielsbacka, Jokela, & Rotkirch, 2017).

Many conflicts between siblings are disputes over property rights (Ross et al., 2011). For example, in an observational study of sibling pairs, studied initially at mean ages of 2.5 and 4.5 years and then followed up 2 years later, the investigators focused on the ways in which such disputes over property were resolved. Possible endings to a conflict were categorized as *compromise*, *reconciliation*, *submission*, and *no resolution*. Most often the conflicts ended without being resolved; compromise or reconciliation was more common when parents did not get involved in the siblings' dispute. Even if conflicts end with or without resolution, siblings have lingering views about their own actions during the dispute, reporting that their own behavior was more positive than that of their sibling (Ross, Smith, Spielmacher, & Recchia, 2004).

These findings suggest that the frequency and quality of conflict is a defining feature of particular family relationships. Sibling conflict may spill over into children's interactions with peers; it also may have long-term consequences for children's social and emotional development. Although only a minority of sibling conflicts are hostile or aggressive (Persram, Della Porta, Scirocco, Howe, & Ross, 2019; Ross, Filyer, Lollis, Perlman, & Martin, 1994), intense, destructive conflict between siblings may foster more serious aggression and behavioral problems in the future (Garcia, Shaw, Winslow, & Yaggi, 2000).

2.2 Conflict resolution

Parents often intervene in toddlers' conflicts over toys, stressing ownership rights and often supporting the peers rather than their own children; however, in one study of Canadian toddlers, mothers were more likely to support their sons than their daughters (Ross, Tesla, Kenyon, & Lollis, 1990). Young children sometimes resolve their own conflicts, without help from adults. For example, in a representative sample of 2-, 3-, and 4-year-old children from eight child care centers in the United States ($N=400$ children), 37% of peer conflicts were resolved by the children themselves (Chen et al., 2001). Even the 2-year-olds resolved 26% of their conflicts, most often by yielding but sometimes by asserting their own will. Across the age groups, the children only sought help from adults on 8% of occasions. In other words, adult caregivers might step in to resolve disputes, but their input was not explicitly sought by the children.

Young children's conflicts can be resolved in several different ways, most simply when one child simply walks away from the dispute. For example, in an intensive study of eight toddlers, four paired with friends and the other four with acquaintances, *disengagement* (turning away or getting distracted) was one of the resolution strategies observed, along with *yielding*, *standing firm*, *bargaining and negotiation*, and *adult intervention* (Ashby & Nielsen-Hewett, 2012). Standing firm against the peer's demands was the most common strategy (seen in half of all conflicts), followed by yielding, with bargaining and negotiation only rarely occurring.

Similar resolution strategies were studied in a sample of 141 firstborn children in Israeli and Palestinian families, observed in their usual child care environment at a mean age of 34 months (Feldman et al., 2010). The conflict resolution strategies identified were *winning* via *the use of force*; *compromise*; *giving in* to the peer's demands; and *turning to adults* for help. Girls were significantly more likely to give in whereas boys were more likely to use force or seek help from adults. The use of force was influenced by gender, not culture, but other means of conflict resolution were influenced by culture and the family environment, including parents' own levels of marital conflict. These findings suggest that some variability in approaches to conflict with peers are influenced by both micro and macro factors in young children's family environments. Thus, while the rate of conflict is not a simple reflection of family adversity or the quality of parent-child relationships (Laible et al., 2008), some family dynamics influence ways in which children pursue and resolve their conflicts with peers.

2.3 Management of peer conflicts in young children's groups

Conflicts with peers are inevitable when young children spend time in groups, in child care centers and preschool classrooms. In some cases, the adults' intervention may expand the conflict to include the intervening adult in the dispute, and thus peer conflict and authority conflict may intertwine within the same social encounter. Such conflicts might be ended swiftly by the adults' intervention. The adults may try to settle the conflict quickly by separating the children or determining who owns the disputed object. Alternatively, the carers or teachers might use the conflict as a 'teaching moment.' They may take the time to mediate the conflict and encourage the children's perspective-taking and social problem-solving (e.g., Gloeckler & Cassell, 2012).

When carers and teachers do intervene in children's conflicts, the dispute may expand the conflict, with the children now resisting or protesting the adult's interfering actions. Although some teachers will see such conflicts as threatening to authority, others see resistance and protest as key elements of living in a democracy which should not be suppressed. For example, in a qualitative study of peer conflicts in Swedish preschool classrooms (Johansson & Emilson, 2016), the investigators reported that "the children expressed resistance by argumentation, distraction, hesitation, ignoring, persistence and by showing courage, solidarity and playfulness" (p. 31). These findings suggest that key social and political skills are already emerging within disputes in preschool classrooms.

In busy child care centers and preschool classrooms, carers and teachers may have only so much time to help children find alternatives to conflict and resolve those conflicts that do occur. However, findings from a study of 3–6-year-olds suggest that, to supplement the efforts of the teachers, robot mediators can help keep the peace in young children's peer groups (Shen, Slovak, & Jung, 2018). The children were observed in pairs in a laboratory playroom, in the presence of an adult experimenter and a small, friendly robot named "Keepon." After an introductory session in which the children became familiar with Keepon's abilities, the children engaged in five play activities, in random order. Keepon facilitated and directed the children's play; for one randomized subsample, the robot also mediated any conflicts that occurred between the peers. In the robot-mediated condition, the majority of children's conflicts (66%) were constructively resolved; in the control condition, when the robot did not play a mediating role, only 18% of conflicts were constructively resolved.

This study draws attention to the role of external mediators in resolving the conflicts that occur in children's peer groups. The findings show that ongoing support from another person throughout a conflict can help young children find ways to resolve their disputes in a positive manner, even if the support is coming from a robot.

3. The interplay between cooperation and conflict

As we have seen, developmental psychologists have long studied very young children's tendencies to engage in cooperation and conflict, but they only rarely study these two types of interactions at the same time. Toddlers' conflicts and cooperative games are hardly ever studied in the same sample or, if both types of interactions are observed, they are hardly ever examined in relation to each other. Similarly, toddlers' prosocial behavior (including cooperation) and aggression have only rarely been studied in relation to each other (see Hay et al., 2021). We thus know less than we might about how very young children balance their conflicts over resources and personal space with their tendencies to engage in mutual play. Rather, studies of cooperation in somewhat older children often took the form of experiments that give children the choice to cooperate or compete, as measured in highly structured experimental games.

3.1 Experimental studies of older children's cooperation vs. competition

Such experiments primarily focused on children of preschool age and older. For example, in the context of social interdependence theory (Deutsch, 1949), experimental studies of children's abilities to cooperate were framed as problem-solving task, a choice between cooperating and competing on the task. Although Deutsch's theory was later expanded to encompass cooperation, competition and conflict (Deutsch, 2014), the initial experiments recruited participants who were of preschool age or older, and measured competition within the parameters of a standardized task, not overt conflict.

For example, Madsen (1967) designed a "Cooperation Board" covered with a piece of paper, with four target circles drawn at different places on the paper. Four children took part in the task at the same time; each could pull on a string to move a pen across the paper. The rules of the game shifted so that the children were required to cooperate (e.g., moving the pen to write across each of the four target circles) or compete to make the pen cross the circles that held their names. In a series of experiments, Madsen observed individual differences in the ability to cooperate but also cultural influences on cooperation and competition (Madsen, 1967; Madsen & Yi, 1975; Shapira & Madsen, 1974), with urban children generally less cooperative than those who grew up in more communal societies.

Another experimental paradigm to measure cooperation featured a toy movie viewer that required two children to operate the apparatus so that a third child could look through an eyepiece (Charlesworth & Dzur, 1987). Girls and boys were equally likely to cooperate on this task, although girls were more likely than boys to use verbal means of achieving that cooperation. Subsequent studies using Madsen's Cooperation Board and/or Charlesworth and Dzur's movie viewer paradigms yielded findings that drew attention to social status and dominance hierarchies in young children's peer groups (for a review, see Green & Rechis, 2006).

Subsequent studies drew attention to parents' support for children's cooperation on such tasks. For example, in one longitudinal study of children tested at 3 and 4 years of age (Ruffman, Slade, Devitt, & Crowe, 2006), the focus was on a mother's conversations with her child while the child was cooperating or competing with a friend on a drawing task; mothers' references to mental states were associated with the child's cooperation/competition scores. More recently, however, the study of cooperation vs. competition has extended beyond those paradigms and instead deployed experimental tasks such as the Prisoner's Dilemma or resource allocation games (e.g., Blake, Rand, Tingley, & Warneken, 2015; Majolo & Maréchal, 2017).

The study of cooperation vs. competition in problem-solving games has provided important information about preschool-aged children's reasoning and social judgments. Parents' reports on preschool children's cooperation and conflict with siblings in relation to the children's temperaments are also informative (Lemery & Goldsmith, 2003). However, the experimental paradigms and questionnaire studies do not provide sufficient information about the relation between young children's cooperation and their spontaneous, emotionally-laden social conflict with siblings or peers.

Our review of both types of interactions in the first 3 years of life highlights the fact that cooperative play and conflict occur naturally when toddlers play with other children. These two types of interaction occur in parallel yet are only rarely studied together. To address this gap in the literature, we conclude this chapter with an illustrative example of the interplay of cooperation and conflict when toddlers interact with unfamiliar peers.

We now present findings from our own longitudinal study of early social development in a sample of firstborn children, observed with unfamiliar peers during a laboratory birthday party.

4. Toddlers' cooperative play and conflict with new acquaintances

4.1 Illustrative findings

4.1.1 The longitudinal study design

The Cardiff Child Development Study, funded by the UK Medical Research Council, had been designed to study the developmental pathway toward emotional and behavioral problems in middle childhood, with assessments at mean ages of 6 months, 1 year, 1.5 years, and 2.5 years, with a follow-up at 7 years (see Hay et al. (2021), for more details.). We measured children's prosocial behavior and other social skills as well as tracing links between early risk factors and later problems. We observed the children directly, with peers as well as parents. In order to observe the children and their parents in a realistic and emotionally challenging social situation, we designed an experimental birthday party paradigm, where the parents and children met other members of the sample in a playroom that was designed to resemble a family living room, decorated with balloons and birthday banners. The party scenario incorporated an emotional challenge in which the toddlers met a costumed character (one of the experimenters dressed as a large teddy bear; for more details see Hay et al., 2016). The challenge was followed by 20 min of free play. The parties were held at the 1- and 2-year assessments; here we report findings from the second birthday party.

Because some families had moved out of the area and others were unable to attend a laboratory session during working hours, only 175 families were able to take part in the birthday party paradigm; others provided questionnaire data only. Three families were scheduled for each laboratory session. If one family canceled, the session was still held with only two families participating. If two families canceled, only individual assessments were carried out and the participating family rescheduled if possible for a later birthday party. The number of children present during the parties varied from two to four.

Video records of the free play session were transcribed using the Peer Interaction Coding System (PICS), which identifies sequences of operationally defined peer-directed actions; key prosocial and forceful actions were recorded with high reliability (see Hay et al. (2021), for more details). The PICS transcripts were coded for episodes of interaction, defined as sequences of peer-directed behavior by two or more participants, separated by intervals of 30 s or more (Hay & Ross, 1982). The episodes of interaction were then examined for evidence of cooperation and conflict. Some longer episodes of interaction contained both cooperation and conflict.

4.1.2 Operational definitions of cooperation and conflict

Cooperation. An episode was judged to contain cooperation if at least two children worked together for at least two moves toward a mutual goal. The goal might be the *literal exchange of objects*; a *playful or pretend exchange*, for example, a tea party in which each child plays a role; or *joint problem-solving*, an occasion where at least two children worked together, performing some instrumental actions on at least one object to meet a goal. The episodes of literal and playful/pretend exchanges required evidence of explicit turn-taking whereby one child acted and then the other reacted, alternating turns throughout the episode. Acting together on an object was not required to feature clear alternation of turns.

An illustrative example of the *literal exchange of objects* would be when one child offered another a piece of plastic birthday cake and the peer reciprocated with a plastic orange. An illustrative episode of *playful/pretend exchange* would be one where at least two children sustained a pretend tea party that featured turn-taking using pretend actions, for example, pouring and drinking invisible tea. Again, such episodes featured alternation of turns. Finally, an example of coordinating instrumental actions on objects would be where two children worked together to push a plastic box filled with balls, the mutual goal being to keep the box moving around the room.

Conflict. An episode of interaction was judged to contain conflict if one child's behavior met with resistance, protest or retaliation from at least one other peer (Hay & Ross, 1982). To meet the criteria for conflict, the episode of interaction must contain at least two peer-directed moves. Thus, instances where one child tugged on a toy and the peer showed no reaction did not meet the criteria for conflict.

An example of *resistance* would be an occasion where one child takes an object from the array of toys a peer is playing with and the peer follows the child who has taken the toy, perhaps reaching for the toy in the other child's hands. Another example would be where one child reaches for or gently contacts an object in the peer's array of toys and the peer withdraws it out of reach. If a child touched a peer who withdrew physically, that was also categorized as resistance.

An example of *protest* would be an episode of interaction where one child takes a toy that the peer had recently put down and the peer fusses or makes a verbal protest, such as "No!" or "Don't!" An example of *retaliation*, which occurred only rarely, would be an occasion where one child tugs a toy away from the peer and the former owner hits the child who now holds the toy.

4.1.3 Cooperative play with new acquaintances

The majority of toddlers who attended the laboratory birthday parties (62%) engaged in cooperative play with their peers. Girls and boys did not differ significantly in their willingness to cooperate with peers; 62% of girls and 61% of boys did so.

However, being paired with toddlers of the same gender seemed to facilitate cooperative play. Most of the groups observed together included both girls and boys; cooperative play was engaged in by 56% of the children in the mixed gender groups.

In contrast, in the same gender groups, 74% of boys ($N=39$) and 100% of girls ($N=12$) cooperated with their new acquaintances.

It is impossible to draw firm conclusions when the number of same-gender groups was so small, but it is possible that the challenges in sustaining cooperative play with peers who are of the opposite gender contribute to the self-initiated preferences for one's own gender that emerge in early childhood (e.g., Shutts, 2015). Alternatively, the growth of understanding of gender as a key feature of the self may affect young children's willingness to spend time with peers of the opposite gender. It would be helpful to study cooperative play and conflict with reference to emerging gender identities.

Group size also mattered. Cooperative play was significantly more likely to occur between pairs of children than in trios, with cooperation shown by 71% of children playing with one other child and 54% of those observed in trios or quartets. These findings suggest that the size of child care groups may influence the likelihood of cooperation between young peers.

4.1.4 Conflict

A majority of the children observed during the laboratory birthday parties (67%) engaged in conflict with their peers. Participation in conflict was not significantly influenced by gender, with 61% of girls and 64% of boys engaging in conflict with their new acquaintances. Nor was the likelihood of conflict significantly influenced by group size or the gender composition of the group.

4.1.5 Cooperation in relation to conflict

Cooperation and conflict were positively correlated. Those children who took part in conflict were also likely to cooperate with their peers (kappa $= 0.19$, $OR = 2.34$, $p < 0.01$), suggesting that both types of interactions represented a willingness to engage socially with unfamiliar peers.

It is important to consider the dynamic balance between cooperation and conflict in the context of the other social and communicative skills developing in the second and third years of life. In this age range, toddlers are developing conversational skills and a broader vocabulary, which means that they can express their desires to peers without necessarily provoking conflict, and they can cooperate in sustaining a dialogue. As children become more proficient with language, their social interactions become verbal as well as physical and their conflicts include verbal arguments.

As an example, a verbal argument between two participants during the laboratory birthday party is presented in Fig. 1. This conversation between

Child A	Child B
	'Little teddy bear'
Makes a screeching noise	
	'Hey, where's the fire? Where's the fire?'
'The fire's over there. That fire's out.'	
(Announces to everyone) 'Look, my water's tipped.'	
	'Your water's tipped.'
'In there…let's put the cups away now. Let's put the cups away now. Let's put the cups away now.'	
'I just need to put the cups away!'	
	'Cups are for tea party, cups are for tea party.'
'I just want to put the cups away.'	
	'No, I just want a tea party.'
(to self) 'Ugh. That was a dirty cake.'	
	'Can I have some, I want a tea party!'

Fig. 1 Conversation and a verbal argument between two toddlers at the laboratory birthday party.

two toddlers illustrates the possibilities for misunderstanding and the cooperative framework needed for such a debate to be kept going over time.

Our example draws attention to some toddlers' abilities to converse and engage in conflict over ideas when speaking to new acquaintances. However, the developmental changes in communicative abilities and interaction skills take place against the background of individual peer relationships that might be more or less likely to facilitate harmonious interaction. Some toddlers may find it especially difficult to balance conflict with cooperation when spending time with various peers. These difficulties might in turn lead to social preferences and social structures in young children's peer relations.

Our findings suggest that it may be especially informative to study conflict and cooperation together in the same samples, in order to learn more about the social dynamics of very young children. These early interaction skills that allow children to balance their own needs with the needs of their peers have implications for children's abilities to form friendships and make successful transitions to social life in school.

5. Conclusions

Our review of the literature on cooperation and conflict in the early years of life has uncovered a bifurcated body of work, where these two fundamental components of social life are only rarely studied in relation to each other. Rather, the two types of social interaction are typically studied within different theoretical frameworks, using different experimental and observational paradigms. At the same time, the findings we have reported from our own longitudinal study of early development suggest that studying these two forms of social engagement in parallel would provide new insights about early social development and children's construction of their social lives.

References

Ashby, N., & Nielsen-Hewett, C. (2012). Approaches to conflict and conflict resolution in toddler relationships. *Journal of Early Childhood Research*, *10*, 145–161. https://doi.org/10.1177/1476718X11430070.

Ashley, J., & Tomasello, M. (1998). Cooperative problem-solving and teaching in preschoolers. *Social Development*, *7*, 143–163.

Blake, P. R., Rand, D. G., Tingley, D., & Warneken, F. (2015). The shadow of the future promotes cooperation in a modified prisoner's dilemma for children. *Scientific Reports*, *5*. article 14559.

Blandon, A. Y., & Scrimgeour, M. B. (2015). Child, parenting, and situational characteristics associated with children's prosocial behavior. *Infant and Child Development*, *24*, 643–660. https://doi.org/10.1002/icd.1910.

Brownell, C., & Carriger, M. S. (1990). Changes in cooperation and self-other differentiation in the second year. *Child Development*, *61*, 1164–1174. https://doi.org/10.1111/j.1467-8624.1990.tb02850.x.

Brownell, C. A., Ramani, G. B., & Zerwas, S. (2006a). Becoming a social partner: Cooperation and social understanding in one- and two-year-olds. *Child Development*, 77(4), 803–821. https://doi.org/10.1111/j.1467-8624.2006.00904.x.

Brownell, C. A., Ramani, G. B., & Zerwas, S. (2006b). Cooperation and social understanding in one- and two-year-olds. *Child Development*, 77(4), 803–821. https://doi.org/10.1111/j.1467-8624.2006.t01-1-.x-i1.

Bruner, J., & Sherwood, V. (1976). Early rule structure: The case of peekaboo. In J. S. Bruner, A. Jolly, & K. Sylva (Eds.), *Play: Its role in evolution and development*. Harmondsworth: Penguin.

Caplan, M., Vespo, J. E., Pedersen, J., & Hay, D. F. (1991). Conflict and its resolution in small groups of one- and two-year-olds. *Child Development*, *62*, 1512–1524. https://www.jstor.org/stable/1130823.

Charlesworth, W. R. (1996). Cooperation and competition: Contributions to an evolutionary and developmental model. *International Journal of Behavioral Development*, *19*, 25–39.

Charlesworth, W. R., & Dzur, C. (1987). Gender comparisons of preschoolers' behavior and resource utilization in group problem solving. *Child Development*, *58*, 191–200. https://www.jstor.org/stable/1130301.

Chen, D. W., Fein, G. G., Killen, M., & Tam, H.-P. (2001). Peer conflicts of preschool children: Issues, resolution, incidence, and age-related patterns. *Early Education and Development*, *12*, 523–544. https://doi.org/10.1207/s15566935eed1204_3.

Deutsch, M. (1949). A theory of co-operation and competition. *Human Relations*, *2*, 129–152. https://psycnet.apa.org/doi/10.1177/001872674900200204.

Deutsch, M. (2014). Cooperation, competition, and conflict. In P. T. Coleman, M. Deutsch, & E. C. Marcus (Eds.), *The handbook of conflict resolution: Theory and practice* (pp. 3–28). Jossey-Bass/Wiley.

Eckerman, C. O., Davis, C. C., & Didow, S. M. (1989). Toddlers' emerging ways of achieving social coordinations with a peer. *Child Development*, *60*, 440–453. https://doi.org/10.2307/1130988.

Eckerman, C. O., & Whitehead, H. (1999). How toddler peers generate coordinated action: A cross-cultural exploration. *Early Education and Development*, *10*, 241–266. https://doi.org/10.1207/s15566935eed1003_2.

Eisenberg, N., Fabes, R. E., & Spinrad, T. (2006). Prosocial development. In W. Damon, R. Lerner, & N. Eisenberg (Eds.), *Social, emotional and personality development: Vol. 3. Handbook of child psychology* (pp. 646–718). New York: Wiley.

Fanger, S. M., Frankel, L. A., & Hazen, N. (2012). Peer exclusion in preschool children's play: Naturalistic observations in a playground setting. *Merrill-Palmer Quarterly*, *58*, 224–254. https://www.jstor.org/stable/23098463.

Fantasia, V., Fasulo, A., Costall, A., & López, B. (2014). Changing the game: Exploring infants' participation in early play routines. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2014.00522.

Feldman, R., Masalha, S., & Derdikman-Eiron, R. (2010). Conflict resolution in the parent–child, marital, and peer contexts and children's aggression in the peer group: A process-oriented cultural perspective. *Developmental Psychology*, *46*, 310–325. https://doi.org/10.1037/a0018286.

Garcia, M. M., Shaw, D., Winslow, E. B., & Yaggi, K. D. (2000). Destructive sibling conflict and the development of conduct problems in young boys. *Developmental Psychology*, *36*, 44–53. https://doi.org/10.1037/0012-1649.36.1.44.

Garvey, C. (1977). *Play*. Cambridge, MA: Harvard University Press.
Gloeckler, L., & Cassell, J. (2012). Teacher practices with toddlers during social problem solving opportunities. *Early Childhood Education*, *40*, 251–257. https://doi.org/10.1007/s10643-011-0495-4.
Goldman, B. D., & Ross, H. S. (1978). Social skills in action: An analysis of early peer games. In J. Glick, & A. Clarke-Stewart (Eds.), *The development of social understanding* (pp. 177–212). Gardner.
Graham, K. L., Burghardt, G. M., & Wiens, J. J. (2010). Current perspectives on the biological study of play: Signs of progress. *Quarterly Review of Biology*, *85*, 393–418. https://doi.org/10.1086/656903.
Green, V. A., & Rechis, R. (2006). Children's cooperative and competitive interactions in limited resource situations: A literature review. *Applied Developmental Psychology*, *27*, 42–59. https://doi.org/10.1016/j.appdev.2005.12.002.
Hay, D. F. (2006). "Yours and mine": Toddlers' talk about possession with familiar peers. *British Journal of Developmental Psychology*, *24*, 39–52. https://doi.org/10.1348/026151005X68880.
Hay, D. F., Castle, J., & Davies, L. (2000). Toddlers' use of force against familiar peers: A precursor of serious aggression? *Child Development*, *71*, 457–467. https://doi.org/10.1111/1467-8624.00157.
Hay, D. F., Mundy, L., Carta, R., Roberts, S., Perra, O., Waters, C., et al. (2011). Known risk factors for violence predict 12-month-olds' aggressiveness with peers. *Psychological Science*, *22*, 1205–1211. https://doi.org/10.1177/0956797611419303.
Hay, D. F., Nash, A., Caplan, M., Swarzentruber, J., Ishikawa, F., & Vespo, J. E. (2011). The emergence of gender differences in physical aggression in the context of conflict between young peers. *British Journal of Developmental Psychology*, *29*, 158–175. https://doi.org/10.1111/j.2044-835X.2011.02028.x.
Hay, D. F., Paine, A. L., Perra, O., Cook, K. V., Hashmi, S., Robinson, C., et al. (2021). Prosocial and aggressive behavior: A longitudinal study. *Monographs of the Society for Research in Child Development*, *86*. Serial No. 341.
Hay, D. F., & Ross, H. S. (1982). The social nature of early conflict. *Child Development*, *53*, 105–113. https://doi.org/10.2307/1129642.
Hay, D. F., van Goozen, S. H. M., Mundy, L., Phillips, R., Roberts, S., Meeuwsen, M., et al. (2016). If you go down to the woods today: Infants distress during a teddy bear's picnic in relation to peer relations and later emotional problems. *Infancy*, *22*(4), 1–19. https://doi.org/10.1111/infa.12172.
Hinde, R. (1979). *Toward understanding relationships*. London: Methuen.
Howe, N., & Recchia, H. (2006). Sibling relations and their impact on children's development. In R. E. Tremblay, R. G. Barr, & R. D. V. Peters (Eds.), *Encyclopedia on early child development* (online publication). http://www.child-encyclopedia.com/documents/Howe-RecchiaANGxp.pdf.
Howe, N., Paine, A., Ross, H., & Recchia, H. (2022). Sibling relations in early and middle childhood. In P. K. Smith, & C. H. Hart (Eds.), *The Handbook of Childhood Social Development* (3rd). Wiley Blackwell.
Hunnius, S., Bekkering, H., & Cillesen, A. H. N. (2009). The association between intention understanding and peer cooperation in toddlers. *European Journal of Developmental Science*, *3*, 368–388.
Johansson, E., & Emilson, A. (2016). Conflicts and resistance: potentials for democracy learning in preschool. *International Journal of Early Years Education*, *24*(1), 19–36.
Johnson, D. W., & Johnson, R. T. (2011). Cooperative learning and social interdependence theory. In R. S. Tindale, et al. (Eds.), *Theory and research on small groups* (pp. 9–33). New York: Kluwer.

Kartner, J., Schumachker, N., & Collard, J. (2014). Socio-cognitive influences on the domain-specificity of prosocial behavior in the second year. *Infant Behavior and Development*, *37*, 665–675. https://doi.org/10.1016/j.infbeh.2014.08.004.

Laible, D., Panfile, T., & Makariev, D. (2008). The quality and frequency of mother-toddler conflict: Links with attachment and temperament. *Child Development*, *79*, 426–443. 0009-3920/2008/7902-0012.

Lemery, K. S., & Goldsmith, H. H. (2003). Genetic and environmental influences on preschool sibling cooperation and conflict: Associations with difficult temperament and parenting style. *Marriage and Family Review*, *33*(1), 75–97. https://doi.org/10.1300/J002v33n01_06.

Lewin, K. (1948). *Resolving social conflicts: Selected papers on group dynamics*. New York: Harper.

Licht, B., Simoni, H., & Perrig-Chello, P. (2008). Conflict between peers in infancy and toddler age: What do they fight about? *Early Years*, *28*, 235–249. https://doi.org/10.1080/09575140802065458.

Madsen, M. C. (1967). Cooperative and competitive motivation of children in three Mexican subcultures. *Psychological Reports*, *20*, 1307–1320.

Madsen, M. C., & Yi, S. (1975). Cooperation and competition in urban and rural children in the republic of South Korea. *International Journal of Psychology*, *10*, 269–274. https://doi.org/10.1080/00207597508247338.

Majolo, B., & Maréchal, L. (2017). Between-group competition elicits within-group cooperation in children. *Scientific Reports*, *7*, 43277.

McHale, S. M., Updegraff, K. A., & Whiteman, S. D. (2012). Sibling relationships and influences in childhood and adolescence. *Journal of Marriage and Family*, *74*, 913–930. https://doi.org/10.1111/j.1741-3737.2012.01011.x.

Meyer, M., Bekkering, H., Haartsen, R., Stapel, J. C., & Hunnius, S. (2015). The role of action prediction and inhibitory control for joint action coordination in toddlers. *Journal of Experimental Child Psychology*, *139*, 203–220. https://doi.org/10.1016/j.jecp.2015.06.005.

NICHD Early Child Care Research Network. (2001). Child care and children's peer interaction at 24 and 36 months: The NICHD study of early child care. *Child Development*, *72*, 1478–1500. https://www.jstor.org/stable/3654400.

Noakes, M. A., & Rinaldi, C. M. (2006). Age and gender differences in peer conflict. *Journal of Youth and Adolescence*, *35*, 881–891. https://doi.org/10.1007/s10964-006-9088-8.

Parten, M. B. (1932). Social participation among pre-school children. *Journal of Abnormal and Social Psychology*, *27*(3), 243–289. https://doi.org/10.1037/h0074524.

Persram, R. J., Della Porta, S., Scirocco, A., Howe, N., & Ross, H. S. (2019). A two-year longitudinal study of naturalistic parent-child-originated and sibling-originated polyadicconflicts. *Merrill-Palmer Quarterly*, *65*, 377–401. https://doi.org/10.13110/merrpalmquar1982.65.4.0377.

Radke-Yarrow, M., Zahn-Waxler, C., & Chapman, M. (1983). Children's prosocial disposition and behavior. In P. Mussen (Ed.), *Handbook of Child Psychology* Wiley.

Reddy, V., & Mireault, G. (2015). Teasing and clowning in infancy. *Current Biology*, *25*, R20–R23. https://doi.org/10.1016/j.cub.2014.09.021.

Ross, H., Conant, C., & Vickar, M. (2011). Property rights and the resolution of social conflict. In H. Ross, & O. Friedman (Eds.), *Vol. 132. Origins of ownership of property. New Directions for Child and Adolescent Development* (pp. 53–64).

Ross, H. S., Filyer, R. E., Lollis, S. P., Perlman, M., & Martin, J. (1994). Administering justice in the family. *Journal of Family Psychology*, *8*, 254–273. https://doi.org/10.1037/0893-3200.8.3.254.

Ross, H. S., & Goldman, B. D. (1977). Establishing new social relations in infancy. In T. Alloway, P. Pliner, & L. Krames (Eds.), *Attachment behaviour: Vol. 3. Advances in the study of communication and affect* (pp. 61–79). New York: Springer.

Ross, H., Smith, J., Spielmacher, C., & Recchia, H. (2004). Shading the truth: Self-serving biases in children's reports of sibling conflicts. *Merrill-Palmer Quarterly*, *50*, 61–85.

Ross, H., Tesla, C., Kenyon, B., & Lollis, S. (1990). Maternal intervention in toddler peer conflict: The socialization of principles of justice. *Developmental Psychology*, *26*, 994–1003. https://doi.org/10.1037/0012-1649.26.6.994.

Rubin, K. H., Burgess, K. B., Dwyer, K. M., & Hastings, P. D. (2003). Predicting preschoolers' externalizing behaviors from toddler temperament, conflict, and maternal negativity. *Developmental Psychology*, *39*, 164–176. https://doi.org/10.1037/0012-1649.39.1.164.

Ruffman, T., Slade, L., Devitt, K., & Crowe, E. (2006). What mothers say and what they do: The relation between parenting, theory of mind, language and conflict/cooperation. *British Journal of Developmental Psychology*, *24*, 105–124.

Schumacher, N., & Kärtner, A. (2015). Explaining interindividual differences in toddlers' collaboration with unfamiliar peers: Individual, dyadic, and social factors. *Frontiers in Psychology*, *6*, 493. https://doi.org/10.3389/fpsyg.2015.00493.

Shantz, C. U. (1987). Conflicts between children. *Child Development*, *58*, 283–305. https://www.jstor.org/stable/1130507.

Shapira, A., & Madsen, M. C. (1974). Cooperative and competitive behavior of kibbutz and urban children in Israel. *Child Development*, *40*, 609–617. https://www.jstor.org/stable/1130507.

Shen, S., Slovak, P., & Jung, M. F. (2018). "Stop. I see a conflict happening." A robot mediator for young children's interpersonal conflict resolution. *Proceedings of the 2018 ACM/IEEE, Chicago, IL, USA*. ACM.

Shutts, K. (2015). Young children's preferences: Gender, race, and social status. *Child Development Perspectives*, *9*, 262–266. https://doi.org/10.1111/cdep.12154.

Tanskanen, A. O., Danielsbacka, M., Jokela, M., & Rotkirch, A. (2017). Sibling conflicts in full and half-sibling households in the UK. *Journal of Biosocial Science*, *49*, 31–47. https://doi.org/10.1017/S0021932016000043.

Tomasello, M., & Vaish, A. (2013). Origins of human cooperation and morality. *Annual Review of Psychology*, *64*, 231–255. https://doi.org/10.1146/annurev-psych-113011-143812.

Tremblay, R. E., Japel, C., Perusse, D., McDuff, P., Boivin, M., Zoccolillo, M., et al. (1999). The search for the age of 'onset' of physical aggression: Rousseau and Bandura revisited. *Criminal Behavior and Mental Health*, *9*, 8–23. https://doi.org/10.1002/cbm.288.

Trevarthen, C., & Reddy, V. (2017). Consciousness in infants. In M. Vellmans, & S. Schneider (Eds.), *Blackwell companion to consciousness* (2nd ed., pp. 45–62). New York: Blackwell.

Verba, M. (1994). The beginnings of collaboration in peer interaction. *Human Development*, *37*, 125–139.

Warneken, F., Chen, F., & Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Development*, 77, 640–653. https://doi.org/10.1111/j.1467-8624.2006.00895.x.

Warneken, F., & Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy*, *11*, 271–294. https://doi.org/10.1111/j.1532-7078.2007.tb00227.x.

CHAPTER NINE

Temporal approaches to the study of friendship: Understanding the developmental significance of friendship change during childhood and adolescence

Julie C. Bowker* **and Jenna Weingarten**
University at Buffalo, The State University of New York, Buffalo, NY, United States
*Corresponding author: e-mail address: jcbowker@buffalo.edu

Contents

Abstract

With longitudinal studies and recently-developed analyses that can model change, it has become increasingly clear that many types of peer experiences during childhood and adolescence are not static. Instead, such experiences change in numerous ways over time, with significant developmental implications. Most of this recent work, however, has focused on change and stability in group-level peer experiences such as peer victimization, peer exclusion, and popularity. As a result, less is known about the extent to which change occurs, and the developmental significance of such changes, in youths' dyadic-level peer experiences and specifically their friendships. Thus, the goal of this chapter is to illustrate the importance of studying friendship change during childhood and adolescence. To do so, we first review the literature on friendship, in general, and then that pertaining to friendship instability and friendship loss, two related areas of friendship change research that have received some theoretical and empirical attention. We introduce a new *friendship development* model, and informed by this

Advances in Child Development and Behavior, Volume 63
ISSN 0065-2407
https://doi.org/10.1016/bs.acdb.2022.04.005

model, encourage peer relations researchers to consider other types of change in friendships. The concluding sections concentrate on the importance of learning more about friendship change for intervention and prevention efforts with youth.

Historically, developmental theory and research adopted a "static" perspective on peer experiences (Bowker & White, 2021). That is, theoretical and empirical attention was paid to the ways in which youth engage and fare with their peers at a single time point. Longitudinal studies, however, began to challenge this view as evidence of significant instability and change in peer experiences began to accumulate. For instance, it is now well-established that significant changes occur in peer experiences over time, such as changes in sociometric status (or the extent to which youth are identified by their peers as highly popular, rejected, neglected, and average; Ollendick, Greene, Francis, & Baum, 1991). In addition, it has become well-documented that certain types of changes in peer experiences, such as increases in peer victimization, contribute to the development of, and exacerbate ongoing social, emotional, and psychological difficulties (Ladd, Ettekal, & Kochenderfer-Ladd, 2017; Troop-Gordon & Ladd, 2005). Further, with the use of new longitudinal statistical methods, including growth curve modeling, it has become clear that some types of peer experiences change, on average, over time, such as peer victimization and peer exclusion, but also that there are distinct groups of children and adolescents who experience varying degrees of stability and change in their peer experiences (e.g., Oncioiu et al., 2020; Shell, Gazelle, & Faldowski, 2014). All of this evidence together has pointed to the importance of moving away from traditional static views of peer experiences and toward more temporal-based approaches that explicitly acknowledge and evaluate the significant change that occurs in children's and adolescents' peer experiences.

Today, it is well-established that change in peer experiences during childhood and adolescence is common and can be developmentally meaningful, but most knowledge pertains to change in *group*-level peer experiences (namely peer victimization, peer exclusion, peer rejection, and popularity), all of which reflect how youth fare with the larger peer group. Considerably less is known about change in *dyadic*-level experiences with friends (Rubin, Bukowski, & Bowker, 2015). Yet, significant change does occur in friendship experiences; for example, approximately 50% of children's and adolescents' friendships are unstable across a single school year (Meter & Card, 2016; Poulin & Chan, 2010). This specific type of friendship change, friendship (in)stability, has been the focus of numerous studies (e.g., Berndt, 2002; Bowker, 2004; Keefe & Berndt, 1996; Poulin & Chan, 2010).

Similar to adult romantic relationships and friendships other types of changes in friendships during childhood and adolescence occur, such as new friendship formation, but are rarely considered (for a study of several types of change in adult friendships, see Jerrome & Wenger, 1999). The goal of this chapter is to illustrate the importance of studying friendship change during childhood and adolescence. To do so, we first review the literature on friendship, in general, and then the work pertaining to friendship instability and friendship loss, two related areas of friendship change research that have received some theoretical and empirical attention. We introduce a new *friendship development* model, and informed by this model, encourage peer relations researchers to consider other types of change in friendships. The concluding sections concentrate on the importance of learning more about friendship change for clinical intervention and prevention efforts with youth, including those suffering due to covid-10 related disruptions to their peer relationships.

1. Friendships during childhood and adolescence

To begin, what is a *friendship*? In the developmental literature, the seemingly straight-forward construct of friendship has been defined and studied in numerous ways. Most current definitions, however, emphasize that it is a close relationship between two peers that involves mutual liking and affection—in other words, two youth who are friends both like and perceive of each other as friends (Rubin et al., 2015). These central features are reflected in the majority of empirical research on friendships in which youth are asked to nominate one or all of their friends, and mutual or reciprocated friendships between two youth are determined (e.g., a mutual friendship would be identified if Zoe nominates Lacey and Lacey nominates Zoe as a friend; Bukowski, Hoza, & Boivin, 1994; Parker & Asher, 1993; Wojslawowicz Bowker, Rubin, Burgess, Booth-LaForce, & Rose-Krasnor, 2006).

Unlike parent-child, sibling, and other familial relationships, friendships are also voluntary relationships in which youth *choose* to be involved. Friendships also tend to be more equalitarian than the relationships that children and adolescents form with the adults in their lives. The other features of friendships during childhood and adolescence can vary across, and within periods of development, but in general, involve intimate disclosure (e.g., talking and sharing secrets and feelings), companionship and recreation (or spending time together hanging out and having fun), instrumental aid and assistance, and a commitment to each other and the relationship (Furman, 1998; Rubin, Fredstrom, & Bowker, 2008; Selman, 1980). Due to the large amount of time spent together, friends also experience more conflict with each other than they do with non-friends or casual acquaintances, but they are also more likely to resolve their conflicts in ways that allow the relationship to continue (Hartup, Laursen, Stewart, & Eastenson, 1988; Newcomb & Bagwell, 1995).

Most children and adolescents (and also adults) have at least one friend (Hartup & Stevens, 1997). In their seminal study, Parker and Asher (1993) reported that almost all of their participants (greater than 99%), who were older children and young adolescents in Grades 3, 4, and 5, nominated at least one same-sex friend (from their school, grade, and classroom), and approximately 78% had at least one mutual friend. In studies of younger children and older adolescents, similar percentages have been reported (e.g., Vaughn, Colvin, Azria, Caya, & Krzysik, 2001). Youth who struggle with peers at the group-level of social complexity, such as socially withdrawn youth who regularly avoid their peers and as a result are rejected, excluded, and victimized, are also able to form at least one mutual friendship during childhood and adolescence (Ladd & Burgess, 1999; Ladd, Kochenderfer-Ladd, Eggum, Kochel, & McConnell, 2011; Rubin, Wojslawowicz, Rose-Krasnor, Booth-LaForce, & Burgess, 2006).

The large majority of youth, including those who struggle with the larger peer group, are able to form friendships because they find peers with whom they are similar (Laursen, 2017). Indeed, it is well-established that youth are attracted to, and subsequently select, peers as friends who are similar to themselves in sex, age, and other demographic characteristics (i.e., race, ethnicity) as well as social reputation (e.g., popularity, peer victimization) and in their display of social behaviors, including the extent to which they engage in aggressive, socially withdrawn, and prosocial behaviors (Bowker et al., 2010; Brechwald & Prinstein, 2011; Haselager, Hartup, van Lieshout, & Riksen-Walraven, 1998). In addition, research consistently shows that not only do youth tend to select and form friendships with similar peers, but also that friends become more similar over time in their behaviors, cognitions, and psychosocial well-being due to socialization processes (e.g., modeling; Giletta et al., 2011; Kandel, 1978a; Prinstein, 2007; Van Zalk, Van Zalk, Kerr, & Stattin, 2011).

In addition to one friend, most youth have multiple friends that form larger networks of peers whom they consider friends and with whom they regularly spend time (e.g., Chan & Poulin, 2007; Dijkstra, Cillessen, & Borch, 2013). Within these larger friendship networks, however, youth typically distinguish between their "best" and "good" or "regular" friends (Bukowski et al., 1994; Parker & Asher, 1993). The distinction between best and good/regular friendships is an important one as *best* friendships tend to be more intimate as well as more influential on the psychosocial and behavioral functioning and well-being of youth (Berndt, 2002; Urberg, 1992; You, Lin, Fu, & Leung, 2013). Consistent with these findings, Sullivan (1953) posited long ago that best friends, or *chums*, best satisfy varying social needs, including those pertaining to intimacy and companionship, and thus carry special importance to the developing child or adolescent. In this regard, one can think of best friends as friend "MVP"s, but growing evidence indicates that larger friendship networks comprising good/regular friends can also serve as important sources of emotional and social support and influence in the lives of youth (Shin, 2017; Sijtsema & Lindenberg, 2018).

These characteristics and features of friendship experiences during childhood and adolescence help to explain the unique influence of friends on social, social-cognitive, and emotional development and psychological well-being (Furman & Rose, 2015; Rubin et al., 2015). For instance, numerous studies indicate that friendships provide a unique context for the development of such social-cognitive skills as perspective-taking skills as well as those involving negotiation and compromise, due to the importance of interactions between relatively equal age-mates for the development of these skills (Piaget, 1932). Friendships also contribute positively to thoughts and feelings about the self and others (as indexed by measures of self-esteem and related constructs), and due to its unique provisions (i.e., offerings of aid and support), friendships can function protectively against the development of loneliness, anxiety, and depressive symptoms (Ladd, Kochenderfer, & Coleman, 1996; Parker & Asher, 1993; Schwartz-Mette, Shankman, Dueweke, Borowski, & Rose, 2020). Most of these studies focused on friendship involvement or whether youth have a mutual friend, but links between friendship quality and psychosocial outcomes have also been revealed. In particular, positive or high-quality friendships (that provide the most support and opportunities for intimacy and companionship) appear to best foster psychological well-being (Buhrmester, 1990; Bukowski et al., 1994; Parker & Asher, 1993).

Although limited in number, several studies show that the unique effects of mutual friendship involvement and friendship quality on adjustment outcomes remain after the effects of group-level peer experiences, such as peer rejection and popularity, are controlled (e.g., Bagwell, Newcomb, & Bukowski, 1998; Ladd, Kochenderfer, & Coleman, 1997).

Significantly, there is also evidence that having mutual friendships, especially when they are characterized by positive qualities, can protect youth from peer difficulties, such as peer victimization, and also promote social successes, such as popularity, with the larger peer group (e.g., Boulton, Trueman, Chau, Whitehand, & Amatya, 1999). For instance, findings from several studies indicate that having mutual friends who are popular can promote adolescents' own popularity due to "basking in the reflected glory" effects (Dijkstra, Cillessen, Lindenberg, & Veenstra, 2010). There is also some evidence that having mutual friendships can decrease some of the detrimental internalizing and externalizing consequences of negative group-level peer experiences, such as peer victimization (e.g., Hodges, Boivin, Vitaro, & Bukowksi, 1999), although a recent meta-analysis indicated that the findings are more mixed in this area than typically described and future research is needed (Schacter et al., 2021). Nevertheless, a growing number of studies show that the positive impact of child and adolescent friendship involvement can persist into young adulthood, which taken together, has led to characterizations of friendship as the *quintessential type of peer experience* during childhood and adolescence (Bagwell et al., 1998; Bagwell, Bowker, & Asher, 2022; Steinhoff & Keller, 2020).

2. Temporal approaches to the study of child and adolescent friendships

Most of the previously described research treated the construct of friendship statically by examining friendship involvement, the characteristics of friends, and the quality of the friendship, as assessed *at a single time point*. In contrast, *temporal* (or dynamic) approaches to the study of peer experiences, in general, do the following: (1) explicitly acknowledge that changes occur in the extent to which youth are involved in different types of peer experiences, in the features of those peer experiences, and how they fare with their peers; (2) involve research designs that allow for changes in peer experiences to be assessed; and (3) evaluate average changes and/or individual changes over time and their developmental impact (Bagwell et al., 2022; Boivin, Petitclerc, Feng, & Barker, 2010; Cillessen & Lansu, 2015). Recent temporal studies of peer victimization, for instance, have shown with short-term longitudinal designs that while peer victimization tends to decline during middle childhood and early adolescence, youth who increase or are stable in their high levels of peer victimization fare worse behaviorally, socially, and psychologically relative to those who decrease or are stably low in the extent to which they are victimized by peers (Brendgen, Girard, Vitaro, Dionne, & Boivin, 2016; Pellegrini & Long, 2002; Troop-Gordon & Ladd, 2005).

Relative to the recent attention paid to change and stability in group-level peer experiences, and especially peer victimization, there has been less theoretical and empirical interest in the changes that occur in youth friendships and their potential developmental significance. This may be, in part, because it is difficult to study certain types of friendship change. For instance, because at any given time point in school settings the large majority of youth have at least one mutual friendship, it is difficult to evaluate large groups of youth who are chronically friendless or who transition from not having any friends (initially friendless) to being later friended. However, some researchers have begun to study new friendship formation (as well as chronic friendlessness and friendship dissolution) when the peer environment changes and there is a re-shuffling of friendships (e.g., the transition from elementary into middle school, in summer camps). This has allowed for the identification of large samples of youth who experience new friendship formation (and other types of change and stability in friendship involvement) over time (e.g., Bowker et al., 2010; Guimond, Laursen, Hartl, & Cillessen, 2019; Hardy, Bukowski, & Sippola, 2002; Ladd, 1990; Parker & Seal, 1996). We encourage investigators interested in friendship involvement change and stability to utilize such designs.

Changes in friendship *quality* can be difficult to study for at least two reasons: (1) ratings of friendship quality are relatively stable across short-periods of time, and also across different friendship partners (e.g., when youth replace one friendship with another); and (2) most youth friendships break-up across short periods of time, making it challenging to study change and stability in quality *within* intact friendships (e.g., Way & Greene, 2006). Shorter intervals between assessments might be helpful in this regard (as suggested by Poulin & Chan, 2010), but we also suspect that change in friendship involvement and quality might not be more commonly studied because it is assumed that when change does occur, it lacks developmental significance, particularly when larger friendship networks exist. We disagree with this assumption and hope that the remaining sections of this chapter will begin to challenge notions pertaining to the significance of friendship change and also provide a new framework for those researchers wishing to pursue this area of inquiry.

2.1 Friendship (in)stability

Research on friendship stability is the largest area of friendship research that has explicitly considered friendship change (or lack thereof; see reviews, Meter & Card, 2016; Poulin & Chan, 2010). This research has been largely based on the idea that many, if not most, friendships during childhood and adolescence come and go due to evolving and developing interests and behaviors of the relationship partners and also poor friendship quality (Laursen, 2017). In addition, it has long been argued that friendships that are *stable* should be best able to satisfy social needs and be the most rewarding and satisfying friendship experiences, thereby best promoting positive growth and well-being (Ladd, 1990; Lessard & Juvonen, 2018; Poulin & Chan, 2010; Savin-Williams & Berndt, 1990). In this line of work, youth typically write or circle (on a roster) the names of their best and good friends at two time points, often across intervals of 6 months to 1 year, and then the reciprocity of the nominations at both time points are determined. Youth with the *same* reciprocated friend at both time points are considered to have a stable friendship (Bowker, 2004; Meter & Card, 2016; Wojslawowicz Bowker et al., 2006), although it is not uncommon for researchers to rely on unilateral nominations of friendship over time so that youth who consistently nominate (or perceive) the same friend over time (regardless of mutuality) are evaluated as having stable friendships (Chan & Poulin, 2007; Lessard & Juvonen, 2018).

Over the past 30 years, there has been consistent evidence showing that friendship *in*stability is a common feature of youth friendship experiences (Meter & Card, 2016). Indeed, many youth, with estimates of approximately 50%, have at least one friendship that is not stable across a single school year (Poulin & Chan, 2010). Importantly, a recent meta-analysis showed that this percentage is similar across studies that: (1) rely on reciprocated versus unilateral nominations of friendship; (2) allow youth to nominate a single versus an unlimited number of friends; (3) compared the stability rates for youth who identify as male and female; and (4) involved children and adolescents with ages ranging from 5 to 18 years (Meter & Card, 2016). There were also no significant differences across study locations, suggesting the relative absence of cross-cultural differences in the prevalence of friendship stability. That said, due to significant cultural differences in the importance of friendships relative to familial relationships and in the meaning of close relationships across cultures, we contend that it may be necessary to conduct culturally-sensitive studies, with quantitative and qualitative assessments of friendship, before strong conclusions are appropriate (Oh et al., 2021; Rubin et al., 2015).

Friendship instability has been consistently associated with a myriad of negative adjustment outcomes. For instance, having at least one unstable friendship over time is related negatively to self-esteem and positive feelings of self-worth, and positively to loneliness and depressive symptoms, during childhood and adolescence (Berndt & Keefe, 1995; Keefe & Berndt, 1996; Ladd, 1990; Parker & Seal, 1996). Having an unstable friendship is also related negatively to academic achievement (Lessard & Juvonen, 2018; Ng-Knight et al., 2019) and higher levels of behavioral problems and peer victimization (Wojslawowicz Bowker et al., 2006).

This is a summary of the literature, however, with at least two important caveats. First, the nature of the impact of having a stable versus unstable friendship likely depends on the characteristics of the friends, such that youth involved in stable friendships with well-adjusted peers appear to fare better than youth who form stable friendships with less well-adjusted youth (Berndt, Hawkins, & Jiao, 1999; Poulin, Dishion, & Haas, 1999). This phenomenon was perhaps best illustrated in a study by Berndt et al. (1999) which showed that across the transition from elementary into middle school, young adolescents increased in behavioral problems when involved in stable friendships with peers high in behavior problems. Thus, for some youth, stable friendships with certain peers may actually lead to individual risk and harm, whereas unstable friendships with such peers may be positive for development. Of course, the characteristics of friends also impacts the degree to which some friendships are maintained over time. Indeed, there is growing indication that friendships between aggressive youth are oftentimes unstable, and also that dissimilarity between friends (in behavioral characteristics and other types of peer experiences) predicts friendship instability, during childhood and adolescence (e.g., Ellis & Zarbatany, 2007; Johnson & Foster, 2005).

Second, although the *unique* developmental significance of *friendship involvement* and *friendship quality* are well-established in studies in which other aspects of friendship adjustment and also group-level peer experiences (e.g., peer rejection) are controlled, the unique contributions of friendship *stability* to the health, development, and well-being during childhood and adolescence are less clear. This is due, in large part, to the findings from several longitudinal studies in which friendship stability was not related significantly to adjustment outcomes when other indicators of friendship adjustment were considered (e.g., Barry & Wentzel, 2006; Barzeva, Richards, Veenstra, Meeus, & Oldehinkel, 2021; Urberg, Değirmencioğlu, & Pilgrim, 1997). These findings might be explained by the strong conceptual and empirical linkages between friendship stability and friendship quality

(Berndt, Hawkins, & Hoyle, 1986; Hiatt, Laursen, Mooney, & Rubin, 2015; Marengo, Rabaglietti, & Tani, 2018; Schneider, Fonzi, Tani, & Tomada, 1997). Youth with positive and high-quality friendships are the most likely to retain them over time. And yet, assessments of friendship quality capture provisions of friendship, such as instrumental aid and offerings of support, which may better explain the benefits of having friends relative to simply having the same friend over time. That said, the findings in this area are mixed, with several studies showing some unique psychosocial benefits of friendship stability (e.g., Ladd et al., 1996; Ng-Knight et al., 2019), and thus we think that this is an important area for future inquiry.

In summary, many youth have friendships that are not stable, and there is some evidence that such instability can lead to psychosocial distress and other adjustment difficulties, likely due to the disruption of friendship provisions (Meter & Card, 2016). The nature of the impact of friendship stability, however, depends, in part, on the characteristics of the friends, and additional research is needed to better understand the unique significance of friendship stability, beyond the positive effects of friendship involvement and positive friendship quality. Two additional future research directions should be mentioned. First, while there is considerable knowledge about the extent to which *single* friendships are stable versus unstable, less is known about the stability of larger friendship networks and the degree to which such stability is uniquely influential on youth adjustment outcomes. One could imagine, however, that considerable change in the composition of youths' larger friendship networks may occur, especially after grade and school transitions, and when it does, could be rather unsettling and disruptive, perhaps even more so relative to the instability of a single friendship (Hardy et al., 2002). To our knowledge, there are only a few studies that simultaneously evaluated the stability of single friendships and the larger friendship network and there were no psychological adjustment outcomes included (e.g., Chan & Poulin, 2007; Değirmencioğlu, Urberg, Tolson, & Richard, 1998), with the exception of Chan and Poulin (2009) who found that the instability of best friendships, but not secondary friendships (which were part of the larger friendship network), was related to depressive mood during early adolescence.

Second, in general, this area of research has a relatively negative perspective on friendship *in*stability, as evident in the types of outcomes studied (e.g., loneliness, poor self-esteem, anxiety, peer victimization). And yet, it is widely acknowledged that some degree of friendship instability is common, and in many cases, a natural consequence of changes in youth (in interests, behaviors, cognitions that leads to dissimilarity between friends; Hartl, Laursen, & Cillessen, 2015; Laursen, 2017) and poor friendship quality. Thus, it may behoove future researchers to include in their temporal studies of friendship more positive theoretically-indicated outcomes (such authenticity) which have been previously linked to static assessments of friendship and well-being (Peets & Hodges, 2018). There is also some indication that youth who change friends (by replacing one friend with another; *different-stable* youth) fare as well as those who are consistently involved in a friendship with the same peer (*same-stable* youth), suggesting future work in this area might also benefit from additional comparisons of *stable*-same and *stable*-different youth (Branje, Frijns, Finkenauer, Engels, & Meeus, 2007; Proulx & Poulin, 2013; Wojslawowicz Bowker et al., 2006), which could help to extend understanding of friendship stability as well as the rarely considered friendship formation and friendship replacement processes.

2.2 Friendship development and friendship loss

Studies of friendship (in)stability should not be confused with studies of friendship loss, as the former area of investigation focuses on consistent relationship involvement with one specific peer over time, while the latter focuses on a specific event in the relationship history of the friendship: its *ending*. In fact, we think it might be helpful to conceptualize friendships during childhood and adolescence as *developing* over time and thereby having relationship histories. This view is similar to how many relationship researchers conceptualize and study friendships and romantic relationships during adulthood as developing and as having histories (with timelines that can be assessed) that include different phases and events, namely the formation and newly-wed/getting-to-know-each-other phase (and the events within this phase), the maintenance phase and its associated events (including enhancement of positive qualities but also increased conflict linked to different events), and then in many cases, the ending or dissolution phase (Harris & Vazire, 2016; Lloyd & Cate, 1985; Sprecher, 2002).

Acknowledging important developmental differences in the relationships that form between adults and between youth (including, but not limited to, greater autonomy and choice in adult relationships), we think that this relationship development and history perspective is important for the study of friendships during childhood and adolescence. See Fig. 1 for our conceptual model of *friendship development* during childhood and adolescence, which we contend can be used to guide questions on how

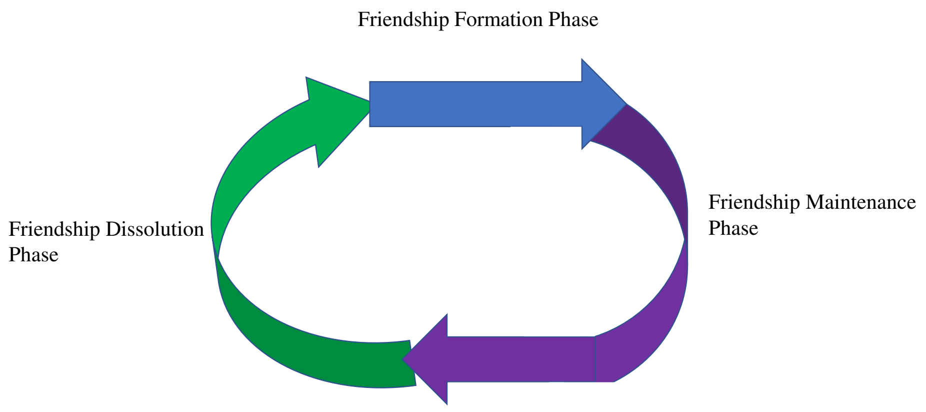


Fig. 1 Friendship development during childhood and adolescence.

friendships, in general, develop (and therefore change) over time, and also individual differences in such relationship development (with such questions as: *Are there individual differences in the speed at which youth proceed through the friendship formation phase, and if so, do such differences have implications for the quality and longevity of the friendship?*). The previously described friendship stability literature fits well in the maintenance and dissolution phases of the model, and the next-to-be discussed friendship loss literature aligns with the dissolution phase. We discuss this temporal model and *friendship development* perspective further in the concluding sections of this chapter, but here emphasize that some conclusions from friendship stability and friendship loss studies are similar. There are also notable differences, which we highlight next.

Although there is a paucity of studies focused specifically on friendship loss, there is some indication that the large majority of youth report at least one recent friendship break-up or dissolution (Bowker, 2011; Flannery & Smith, 2021; Ford, Cllishaw, Meltzer, & Goodman, 2007). Of course, these findings are similar to those in the friendship stability literature, but the percentages of youth reporting friendship losses (e.g., 86%, Flannery & Smith, 2021) tend to be greater than those identified as having unstable friendships (~50%, Meter & Card, 2016). This is likely due to important methodological differences. Friendship loss is typically evaluated with self-reported losses (regardless of the specific friend; Bowker, 2011; Flannery & Smith, 2021), unlike friendship stability, which is usually determined by mutual or unilateral assessments of same-sex and same-school/same-grade friends. Thus, it may be that when directly asked about dissolutions without the identity restrictions associated with friendship nominations, youth are able to more fully report on the friendship dissolutions that they have experienced, including those with friends outside of their immediate school, but whose loss was still meaningful.

Similar to the findings on friendship stability, the results from several studies suggest that the experience of losing a friend can function as a significant interpersonal stressor and challenge. As one example, Ford et al. (2007) found that "permanently falling out with a best friend" during late childhood and early adolescence predicted subsequent psychopathology symptoms, above and beyond the significant effects of recent parental separation, parental mental health problems, and child physical health problems. One study also positively linked the number of dissolutions to depressive symptoms during early adolescence (Flannery & Smith, 2021). In addition, Benenson and Christakos (2003) found that most young adolescents reported feeling badly when imagining hypothetical friendship losses.

In some ways, these results—primarily from studies involving older children and young adolescents—parallel those showing that romantic relationship dissolution can increase psychosocial risk for older adolescents and young adults (Monroe, Rohde, Seeley, & Lewinsohn, 1999). But, since friendships are the most important peer relationship during childhood and early adolescence, it may be that older children and young adolescents are not yet well-prepared to cope with the loss of a friendship, just as older adolescents are thought to be ill-prepared to cope with the loss of a romantic relationship. Studies with older adolescents, however, will be necessary to evaluate this developmental timing hypothesis as it applies to friendship loss.

There is, however, some indication that the *type* of dissolution might matter. In a series of studies, Bowker and colleagues (Bowker, 2011; Bowker, Etkin, & Dirks, 2022; Bowker, White, & Weingarten, 2022) made the distinction between *complete* and *downgrade* dissolutions. The overwhelming majority of research on dissolution (and also friendship stability) has focused exclusively on *complete dissolutions*, or when best friendships end and the two former best friends no longer consider each other friends (Bowker, 2011; Bowker et al., 2010). But some youth may experience *downgrade* dissolutions wherein a best friendship dissolves, but the two former best friends remain good or close friends with a lesser degree of intimacy, stability, and interpersonal influence (Urberg, 1992). To make this distinction, Bowker et al. developed specific questions for youth to report on complete dissolutions ("In the last three months, have you had a best friend of the same-sex with whom you are now no longer friends?") and downgrade dissolutions ("In the last three months, have you had a best friend of the same-sex with whom you are now only a good friend?"; Bowker, Etkin, & Dirks, 2022). In addition, Bowker et al. have posited that changes from best to close friendships might also be experienced by children and young adolescents as interpersonal losses—losses that are not "complete" but that remain psychologically painful due to the interpersonal rejection.

Initial evidence supported the notion that downgrades occur and may even be more common than complete dissolutions (55% vs 36%; Bowker, 2011). Findings from Bowker (2011) also suggested that downgrades are psychologically stressful, especially for youth who are unable to form new best friendships. In this initial pilot study ($N=77$), young adolescents who experienced downgrade dissolutions only reported elevated levels of loneliness relative to those who did not experience downgrade dissolutions when they did not have a current best friendship. Most previous research on friendship stability and friendship loss has been limited by a neglect of new friendship replacement, but the Bowker (2011) findings suggest that forming a new best friendship may function protectively. However, Lessard and Juvonen (2018) found that friendship gains interfered with academic achievement, similar to friendship instability. Thus, we encourage researchers to further consider whether friendship gains (and new friendship replacement) may have both positive and negative adjustment trade-offs, perhaps due to the offerings of support provided by the new relationships, but also the energy and effort needed to establish new relationships. Nevertheless, findings from the recent *UB Changes in Friendship Project,* which involved a larger sample (with $N=271$) relative to the Bowker (2011) study and four assessments of friendship dissolution across 3 month intervals (see Bowker, Adams, Bowker, Fisher, & Spencer, 2016; Markovic & Bowker, 2017), have been consistent with Bowker (2011) in their indication that: (1) young adolescents think about complete and downgrade dissolutions differently, and that such cognitions can help to explain the ways in which the two types of dissolutions are related to psychological and social adjustment outcomes (Bowker, Etkin, & Dirks, 2022); and (2) the characteristics surrounding complete and downgrade dissolutions, such as the emotional reactions that you report following such dissolutions, can also help to explain when such dissolutions increase psychological risk (Bowker & White, 2021).

3. Future directions

We hope that we have been successful in convincing the reader that the research on friendship stability and friendship loss illustrates some of the benefits of utilizing temporal approaches in the study of friendships. From these two areas of research, it has become clear that friendships break-up, perhaps in a variety of different ways, and that such break-ups can lead to significant adjustment implications for the developing child and adolescent. Returning to the newly-proposed model of *friendship development*, however, we contend that other types of friendship change should be more carefully evaluated, such as friendship formation, and a relationship history or development perspective is needed to further expand our understanding of the ways in which friendships develop over time, and when and why friendship development might become problematic for the relationship itself and also for its two relationship members. We next describe research questions and study designs to optimize knowledge about friendship formation, and then we discuss the need to more carefully consider the role of new media in the friendship development process.

3.1 Friendship formation

As noted previously, *friendship formation* is a type of friendship change that can be difficult to capture and evaluate. Large samples and creative research designs, however, can overcome this challenge. So too can experimental designs in which unfamiliar youth are brought together for repeated interactions to see which youth "hit it off" and desire to be friends (see Gottman & Graziano, 1983). Additional knowledge about friendship formation is needed to further understand different types of friendship change and also for clinical intervention efforts with youth who are chronically friendless, and as a result, at considerable risk for short- and long-term psychological and social difficulties (Troop-Gordon & Ladd, 2005; Wojslawowicz Bowker et al., 2006).

What we currently know about friendship formation is that similarities in a myriad of different ways, including risk taking, drug and alcohol use, behaviors, psychological distress, and academic achievement, appear to facilitate the friendship formation process (Kandel, 1978b; Kupersmidt, DeRosier, & Patterson, 1995; Laursen, 2017). However, the reasons for these findings and the exact processes involved are not known. For instance, it is clear that similarities lead to attraction and perhaps draw previously unfamiliar or unacquainted youth together into the same physical space (and proximity is another known predictor of friendship formation), but what happens next? What can we teach youth who are struggling to form friends—just to find a similar peer? Perhaps find a similar peer and then also behave in socially competent ways, as there is also evidence showing that socially competent youth are more likely to form friends relative to those with poor social skills (Rubin et al., 2015)? But more knowledge is needed here such as whether the similarities themselves lead to shared activities, which in turn lead to friendship formation? Perhaps the similarities also lead to fun and intimate disclosure, and these are the "special ingredients" by which similarities promote friendship formation? The observational research by Gottman and Graziano (1983) suggests that this might be the case, but this work was done with children ages 3–9 years in a laboratory context, and the extent to which the social processes (e.g., effective and clear communication) which led youth to "hit it off" are generalizable to other ages of youth and other settings is not known.

In further study of friendship formation, we also encourage researchers to consider friendship formation as a phase rather than a singular event. Similar to how this phase is conceptualized with adults and their friendships and romantic relationships, it is likely that youth spend some time getting to know each other prior to viewing each other as friends, and once friends, there is a newly-wed like period when the relationship is not well-established and when conflict is low and positive interactions are common. To our knowledge, this is a novel way to conceptualize the ways in which child and adolescent friendships form, but consistent with our *friendship development* conceptual model and the widely-held notion that many types of relationships take time to develop (Hinde, 1997). Also consistent with our model but not yet studied is the strong possibility that some friendships during childhood and adolescence could dissolve but then re-form and become friendships again (on-again, off-again friendships), perhaps with a shortened or more uncertain friendship formation phase. In general, additional knowledge about what tends to work and not work during the friendship formation phase to bring youth together into mutually committed relationships could provide important new tools and clues that could be taught, by clinicians and school counselors, to youth who are struggling to form friendships.

3.2 New social media

In future research on friendship change, we encourage investigators to consider the contributions of both off-line and on-line interactions, behaviors, and experiences. To our knowledge, there has not been a single study relating online interactions to friendship stability, but the rise of online tools and platforms and new social media use, particularly among adolescents, has fundamentally changed the nature of adolescent peer relationships. For instance, Nesi, Choukas-Bradley, and Prinstein (2018) describe changes in frequency of contact among friends (offering both rapid social support and various forms of negative feedback), increased friendship demands (including expectations to "like" or comment on a friend's post), an altered communication style (potentially leading to changes in the quality of interactions), and new opportunities to virtually meet and communicate with geographically-distant peers.

Prior research has explored the significance of new media to (offline) friendship quality, and even so, the findings have been unclear. A recent study of the impact of social media use on friendship closeness showed both positive and negative impacts of social media use on friendship closeness

(Pouwels, Valkenburg, Beyens, van Driel, & Keijsers, 2021). In a similar vein, a review of the literature indicated that social media use can lead to both perceived social support and social isolation (Best, Manktelow, & Taylor, 2014). These mixed findings suggest that a number of factors—including patterns of personal use and the social media platform—may impact friendship change and stability in different ways. It is also possible that new social media serves to magnify friendship quality, such that strong and mutual friendships may flourish, while some friendships may struggle under the expected high frequency of contact and increased friendship demands associated with new social media use. Further work should probe the conditions under which new social media use is predictive of friendship change, including friendship formation, stability and maintenance, and loss, which in turn, could lead to greater knowledge to not only help adolescents navigate their online peer interactions and relationships, but also those wishing to form, enhance, or even end their poor quality or otherwise unsatisfying friendships.

4. Conclusions

We completed this chapter as the world enters its third year of the covid-19 pandemic. Although significant and remarkable improvements in virus prevention and treatment have occurred, the world's youth population, along with their parents, continues to struggle due to numerous uncertainties, and also changes, including the disruption of their schooling and close friendships. Indeed, in this third year, many youth are continuing to take classes remotely from home, and those who are back in school continue to be forced into periodic quarantine away from their friends and peers for extended periods of time.

While few studies on the impact of covid-19 and its stressors on youth friendships exist (for several recent exceptions, see Bernasco, Nelemans, van der Graaff, & Branje, 2021; Gadassi Polack et al., 2021; Juvonen, Schacter, & Lessard, 2021), it is becoming increasingly clear that youth today are struggling psychologically—more so than prior to the start of the pandemic (Gruber et al., 2021; Magson et al., 2021). While there are many factors that can account for this psychological distress, the early research in this area suggests that the unwanted disruptions to youth friendships is one likely contributor. Long before the pandemic, it was well-established that friendships during childhood and adolescence tend to break up over time. What still needs to be known, however, is how and why friendships develop in the first place, other contributors to friendship maintenance (besides similarity), and when and why friendship dissolution can lead to positive versus negative adjustment trade-offs. Such knowledge would be useful for theory and research on friendship and friendship development. It could also be useful for clinical intervention and prevention efforts aimed at helping youth form initial friendships and renew friendships after summer breaks and grade transitions – and perhaps also after quarantine periods and other school disruptions that have occurred repeatedly over the past 3 years in all parts of the globe.

References

Bagwell, C. L., Bowker, J. C., & Asher, S. A. (2022). *Back to the dyad: Future directions for friendship research*. Merrill-Palmer Quarterly.

Bagwell, C. L., Newcomb, A. F., & Bukowski, W. M. (1998). Preadolescent friendship and peer rejection as predictors of adult adjustment. *Child Development*, *69*, 140–153.

Barry, C. M., & Wentzel, K. R. (2006). Friend influence on prosocial behavior: The role of motivational factors and friendship characteristics. *Developmental Psychology*, *42*(1), 153.

Barzeva, S. A., Richards, J. S., Veenstra, R., Meeus, W. H., & Oldehinkel, A. J. (2021). Quality over quantity: A transactional model of social withdrawal and friendship development in late adolescence. *Social Development*, 1–21.

Benenson, J. F., & Christakos, A. (2003). The greater fragility of females' versus males' closest same-sex friendships. *Child Development*, *74*, 1123–1129.

Bernasco, E. L., Nelemans, S. A., van der Graaff, J., & Branje, S. (2021). Friend support and internalizing symptoms in early adolescence during COVID-19. *Journal of Research on Adolescence*, *31*(3), 692–702.

Berndt, T. J. (2002). Friendship quality and social development. *Current Directions in Psychological Science*, *11*(1), 7–10.

Berndt, T. J., Hawkins, J. A., & Hoyle, S. G. (1986). Changes in friendship during a school year: Effects on children's and adolescents' impressions of friendship and sharing with friends. *Child Development*, *57*, 1284–1297.

Berndt, T. J., Hawkins, J. A., & Jiao, Z. (1999). Influences of friends and friendships on adjustment to junior high school. *Merrill-Palmer Quarterly*, *45*, 13–41.

Berndt, T. J., & Keefe, K. (1995). Friends' influence on adolescents' adjustment to school. *Child Development*, *66*(5), 1312–1329.

Best, P., Manktelow, R., & Taylor, B. (2014). Online communication, social media and adolescent wellbeing: A systematic narrative review. *Children and Youth Services Review*, *41*, 27–36.

Boivin, M., Petitclerc, A., Feng, B., & Barker, E. D. (2010). The developmental trajectories of peer victimization in middle to late childhood and the changing nature of their behavioral correlates. *Merrill-Palmer Quarterly (1982)*, 231–260.

Boulton, M. J., Trueman, M., Chau, C. A. M., Whitehand, C., & Amatya, K. (1999). Concurrent and longitudinal links between friendship and peer victimization: Implications for befriending interventions. *Journal of Adolescence*, *22*(4), 461–466.

Bowker, A. (2004). Predicting friendship stability during early adolescence. *Journal of Early Adolescence*, *24*, 85–112.

Bowker, J. C. (2011). Examining two types of best friendship dissolution during early adolescence. *The Journal of Early Adolescence*, *31*(5), 656–670.

Bowker, J. C., Adams, R. E., Bowker, M. H., Fisher, C., & Spencer, S. V. (2016). Same-and other-sex popularity and preference during early adolescence. *The Journal of Early Adolescence*, *36*(5), 704–722.

Bowker, J. C., & White, H. I. (2021). Studying peers in research on social withdrawal: Why broader assessments of peers are needed. Child Development Perspectives, 15(2), 90–95.
Bowker, J. C., Etkin, R. E., & Dirks, M. A. (2022). When best friendships end: Young adolescents responses to best friendship dissolution. *Journal of Research on Adolescence*.
Bowker, J. C., Fredstrom, B. K., Rubin, K. H., Rose-Krasnor, L., Booth-LaForce, C., & Laursen, B. (2010). Distinguishing children who form new best-friendships from those who do not. *Journal of Social and Personal Relationships*, *27*(6), 707–725.
Bowker, J. C., & White, H. I. (2021). Studying peers in research on social withdrawal: Why broader assessments of peers are needed. *Child Development Perspectives*, *15*(2), 90–95.
Bowker, J. C., White, H. I., & Weingarten, J. (2022). *An exploratory study of emotional responses to friendship dissolution*. (Manuscript under review).
Branje, S. J., Frijns, T. O. M., Finkenauer, C., Engels, R., & Meeus, W. I. M. (2007). You are my best friend: Commitment and stability in adolescents' same-sex friendships. *Personal Relationships*, *14*(4), 587–603.
Brechwald, W. A., & Prinstein, M. J. (2011). Beyond homophily: A decade of advances in understanding peer influence processes. *Journal of Research on Adolescence*, *21*(1), 166–179.
Brendgen, M., Girard, A., Vitaro, F., Dionne, G., & Boivin, M. (2016). Personal and familial predictors of peer victimization trajectories from primary to secondary school. *Developmental Psychology*, *52*(7), 1103–1114. https://doi.org/10.1037/dev0000107.
Buhrmester, D. (1990). Intimacy of friendship, interpersonal competence, and adjustment during preadolescence and adolescence. *Child Development*, *61*(4), 1101–1111.
Bukowski, W. M., Hoza, B., & Boivin, M. (1994). Measuring friendship quality during pre- and early adolescence: The development and psychometric properties of the friendship qualities scale. *Journal of Social and Personal Relationships*, *11*, 471–484.
Chan, A., & Poulin, F. (2007). Monthly changes in the composition of friendship networks in early adolescence. *Merrill-Palmer Quarterly*, *53*, 578–602.
Chan, A., & Poulin, F. (2009). Monthly instability in early adolescent friendship networks and depressive symptoms. *Social Development*, *18*(1), 1–23.
Cillessen, A. H., & Lansu, T. A. (2015). Stability, correlates, and time-covarying associations of peer victimization from grade 4 to 12. *Journal of Clinical Child & Adolescent Psychology*, *44*(3), 456–470.
Değirmencioğlu, S. M., Urberg, K. A., Tolson, J. M., & Richard, P. (1998). Adolescent friendship networks: Continuity and change over the school year. *Merrill-Palmer Quarterly (1982)*, 313–337.
Dijkstra, J. K., Cillessen, A. H., & Borch, C. (2013). Popularity and adolescent friendship networks: Selection and influence dynamics. *Developmental Psychology*, *49*(7), 1242.
Dijkstra, J. K., Cillessen, A. H., Lindenberg, S., & Veenstra, R. (2010). Basking in reflected glory and its limits: Why adolescents hang out with popular peers. *Journal of Research on Adolescence*, *20*(4), 942–958.
Ellis, W., & Zarbatany, L. (2007). Explaining friendship formation and friendship stability: The role of children's and friends' aggression and victimization. *Merrill-Palmer Quarterly*, *53*(1), 79–104. https://doi.org/10.1353/mpq.2007.0001.
Flannery, K. M., & Smith, R. L. (2021). Breaking up (with a friend) is hard to do: An examination of friendship dissolution among early adolescents. *The Journal of Early Adolescence*, *41*(9), 1368–1393.
Ford, T., Cllishaw, S., Meltzer, H., & Goodman, R. (2007). A prospective study of childhood psychopathology: Independent predictors of change over three years. *Social Psychiatry and Psychiatric Epidemiology*, *42*, 953–961.
Furman, W. (1998). The measurement of friendship perceptions: Conceptual and methodological issues. In W. M. Bukowski, A. F. Newcomb, & W. W. Hartup (Eds.), *The company they keep: Friendship in childhood and adolescence* (pp. 41–65). Cambridge University Press.

Furman, W., & Rose, A. J. (2015). Friendships, romantic relationships, and peer relationships. In M. E. Lamb, & R. M. Lerner (Eds.), *Handbook of child psychology and developmental science: Socioemotional processes* (pp. 932–974). John Wiley & Sons, Inc.

Gadassi Polack, R., Sened, H., Aubé, S., Zhang, A., Joormann, J., & Kober, H. (2021). Connections during crisis: Adolescents' social dynamics and mental health during COVID-19. *Developmental Psychology*, *57*(10), 1633–1647. https://doi.org/10.1037/dev0001211.

Giletta, M., Scholte, R. H. J., Burk, W. J., Engels, R. C. M. E., Larsen, J. K., Prinstein, M. J., et al. (2011). Similarity in depressive symptoms in adolescents' friendship dyads: Selection or socialization? *Developmental Psychology*, *47*(6), 1804–1814. https://doi.org/10.1037/a0023872.

Gottman, J. M., & Graziano, W. G. (1983). How children become friends. *Monographs of the Society for Research in Child Development*, *48*(3), 1–86.

Gruber, J., Prinstein, M. J., Clark, L. A., Rottenberg, J., Abramowitz, J. S., Albano, A. M., et al. (2021). Mental health and clinical psychological science in the time of COVID-19: Challenges, opportunities, and a call to action. *American Psychologist*, *76*(3), 409.

Guimond, F. A., Laursen, B., Hartl, A. C., & Cillessen, A. H. (2019). Differences in internalizing symptoms anticipate adolescent friendship dissolution. *Journal of Research on Adolescence*, *29*(4), 924–937.

Hardy, C. L., Bukowski, W. M., & Sippola, L. K. (2002). Stability and change in peer relationships during the transition to middle-level school. *Journal of Early Adolescence*, *22*(2), 117–142. https://doi.org/10.1177/0272431602022002001.

Harris, K., & Vazire, S. (2016). On friendship development and the Big Five personality traits. *Social and Personality Psychology Compass*, *10*(11), 647–667.

Hartl, A. C., Laursen, B., & Cillessen, A. H. (2015). A survival analysis of adolescent friendships: The downside of dissimilarity. *Psychological Science*, *26*(8), 1304–1315.

Hartup, W. W., Laursen, B., Stewart, M. I., & Eastenson, A. (1988). Conflict and the friendship relations of young children. *Child Development*, 1590–1600.

Hartup, W. W., & Stevens, N. (1997). Friendships and adaptation in the life course. *Psychological Bulletin*, *121*(3), 355–370.

Haselager, G. J., Hartup, W. W., van Lieshout, C. F., & Riksen-Walraven, J. M. A. (1998). Similarities between friends and nonfriends in middle childhood. *Child Development*, *69*(4), 1198–1208.

Hiatt, C., Laursen, B., Mooney, K. S., & Rubin, K. H. (2015). Forms of friendship: A person-centered assessment of the quality, stability, and outcomes of different types of adolescent friends. *Personality and Individual Differences*, 77, 149–155.

Hinde, R. A. (1997). *Relationships: A dialectical perspective*. Psychology Press.

Hodges, E. V. E., Boivin, M., Vitaro, F., & Bukowksi, W. M. (1999). The power of friendship protection against an escalating cycle of peer victimization. *Developmental Psychology*, *35*, 94–101.

Jerrome, D., & Wenger, G. C. (1999). Stability and change in late-life friendships. *Ageing & Society*, *19*(6), 661–676.

Johnson, D. R., & Foster, S. L. (2005). The relationship between relational aggression in kindergarten children and friendship stability, mutuality, and peer liking. *Early Education & Development*, *16*(2), 141–160.

Juvonen, J., Schacter, H. L., & Lessard, L. M. (2021). Connecting electronically with friends to cope with isolation during COVID-19 pandemic. *Journal of Social and Personal Relationships*, *38*(6), 1782–1799.

Kandel, D. B. (1978a). Homophily, selection, and socialization in adolescent friendships. *American Journal of Sociology*, *84*(2), 427–436.

Kandel, D. B. (1978b). Similarity in real-life adolescent friendship pairs. *Journal of Personality and Social Psychology*, *36*(3), 306–312.

Keefe, K., & Berndt, T. J. (1996). Relations of friendship quality to self-esteem in early adolescence. *The Journal of Early Adolescence*, *16*(1), 110–129.
Kupersmidt, J. B., DeRosier, M. E., & Patterson, C. P. (1995). Similarity as the basis for children's friendships: The roles of sociometric status, aggressive and withdrawn behavior, academic achievement and demographic characteristics. *Journal of Social and Personal Relationships*, *12*(3), 439–452.
Ladd, G. W. (1990). Having friends, keeping friends, making friends, and being liked by peers in the classroom: Predictors of children's early school adjustment? *Child Development*, *61*, 1081–1100.
Ladd, G. W., & Burgess, K. B. (1999). Charting the relationship trajectories of aggressive, withdrawn, and aggressive/withdrawn children during early grade school. *Child Development*, *70*(4), 910–929.
Ladd, G. W., Ettekal, I., & Kochenderfer-Ladd, B. (2017). Peer victimization trajectories from kindergarten through high school: Differential pathways for children's school engagement and achievement? *Journal of Educational Psychology*, *109*(6), 826–841. https://doi.org/10.1037/edu0000177.
Ladd, G. W., Kochenderfer, B. J., & Coleman, C. C. (1996). Friendship quality as a predictor of young children's early school adjustment. *Child Development*, *67*(3), 1103–1118.
Ladd, G. W., Kochenderfer, B. J., & Coleman, C. C. (1997). Classroom peer acceptance, friendship, and victimization: Distinct relational systems that contribute uniquely to children's school adjustment? *Child Development*, 1181–1197.
Ladd, G. W., Kochenderfer-Ladd, B., Eggum, N. D., Kochel, K. P., & McConnell, E. M. (2011). Characterizing and comparing the friendships of anxious-solitary and unsociable preadolescents. *Child Development*, *82*(5), 1434–1453.
Laursen, B. (2017). Making and keeping friends: The importance of being similar. *Child Development Perspectives*, *11*(4), 282–289.
Lessard, L. M., & Juvonen, J. (2018). Losing and gaining friends: Does friendship instability compromise academic functioning in middle school? *Journal of School Psychology*, *69*, 143–153.
Lloyd, S. A., & Cate, R. M. (1985). Attributions associated with significant turning points in premarital relationship development and dissolution. *Journal of Social and Personal Relationships*, *2*(4), 419–436.
Magson, N. R., Freeman, J. Y., Rapee, R. M., Richardson, C. E., Oar, E. L., & Fardouly, J. (2021). Risk and protective factors for prospective changes in adolescent mental health during the COVID-19 pandemic. *Journal of Youth and Adolescence*, *50*(1), 44–57.
Marengo, D., Rabaglietti, E., & Tani, F. (2018). Internalizing symptoms and friendship stability: Longitudinal actor-partner effects in early adolescent best friend dyads. *The Journal of Early Adolescence*, *38*(7), 947–965.
Markovic, A., & Bowker, J. C. (2017). Friends also matter: Examining friendship adjustment indices as moderators of anxious-withdrawal and trajectories of change in psychological maladjustment. *Developmental Psychology*, *53*(8), 1462.
Meter, D. J., & Card, N. A. (2016). Stability of children's and adolescents' friendships: A meta-analytic review. *Merrill-Palmer Quarterly*, *62*(3), 252–284.
Monroe, S. M., Rohde, P., Seeley, J. R., & Lewinsohn, P. M. (1999). Life events and depression in adolescence: Relationship loss as a prospective risk factor for first onset of major depressive disorder. *Journal of Abnormal Psychology*, *108*(4), 606–614. https://doi.org/10.1037/0021-843X.108.4.606.
Nesi, J., Choukas-Bradley, S., & Prinstein, M. J. (2018). Transformation of adolescent peer relations in the social media context: Part 1—A theoretical framework and application to dyadic peer relationships. *Clinical Child and Family Psychology Review*, *21*(3), 267–294.

Newcomb, A. F., & Bagwell, C. L. (1995). Children's friendship relations: A meta-analytic review. *Psychological Bulletin*, *117*(2), 306–347. https://doi.org/10.1037/0033-2909.117.2.306.

Ng-Knight, T., Shelton, K. H., Riglin, L., Frederickson, N., McManus, I. C., & Rice, F. (2019). 'Best friends forever'? Friendship stability across school transition and associations with mental health and educational attainment. *British Journal of Educational Psychology*, *89*(4), 585–599.

Oh, W., Bowker, J. C., Santos, A. J., Ribeiro, O., Guedes, M., Freitas, M., et al. (2021). Distinct profiles of relationships with mothers, fathers, and best friends and social-behavioral functioning in early adolescence: A cross-cultural study. *Child Development*, *92*(6), e1154–e1170.

Ollendick, T. H., Greene, R. W., Francis, G., & Baum, C. G. (1991). Sociometric status: Its stability and validity among neglected, rejected and popular children. *Journal of Child Psychology and Psychiatry*, *32*(3), 525–534.

Oncioiu, S. I., Orri, M., Boivin, M., Geoffroy, M. C., Arseneault, L., Brendgen, M., et al. (2020). Early childhood factors associated with peer victimization trajectories from 6 to 17 years of age. *Pediatrics*, *145*(5).

Parker, J. G., & Asher, S. R. (1993). Friendship and friendship quality in middle childhood: Links with peer group acceptance and feelings of loneliness and social dissatisfaction. *Developmental Psychology*, *29*, 611–621.

Parker, J. G., & Seal, J. (1996). Forming, losing, renewing, and replacing friendships: Applying temporal parameters to the assessment of children's friendship experiences. *Child Development*, *67*, 2248–2268.

Peets, K., & Hodges, E. V. (2018). Authenticity in friendships and well-being in adolescence. *Social Development*, *27*(1), 140–153.

Pellegrini, A. D., & Long, J. D. (2002). A longitudinal study of bullying, dominance, and victimization during the transition from primary school through secondary school. *British Journal of Developmental Psychology*, *20*(2), 259–280.

Piaget, J. (1932). *The moral judgment of the child*. London: Routledge & Kegan Paul.

Poulin, F., & Chan, A. (2010). Friendship stability and change in childhood and adolescence. *Developmental Review*, *30*(3), 257–272. https://doi.org/10.1016/j.dr.2009.01.001.

Poulin, F., Dishion, T. J., & Haas, E. (1999). The peer influence paradox: Friendship quality and deviancy training within male adolescent friendships. *Merrill-Palmer Quarterly*, *45*(1), 42–61.

Pouwels, J. L., Valkenburg, P. M., Beyens, I., van Driel, I. I., & Keijsers, L. (2021). Social media use and friendship closeness in adolescents' daily lives: An experience sampling study. *Developmental Psychology*, *57*(2), 309.

Prinstein, M. J. (2007). Moderators of peer contagion: A longitudinal examination of depression socialization between adolescents and their best friends. *Journal of Clinical Child and Adolescent Psychology*, *36*(2), 159–170.

Proulx, M. F., & Poulin, F. (2013). Stability and change in kindergartners' friendships: Examination of links with social functioning. *Social Development*, *22*(1), 111–125.

Rubin, K. H., Bukowski, W. M., & Bowker, J. C. (2015). Children in peer groups. In M. H. Bornstein, T. Leventhal, & R. M. Lerner (Eds.), *Handbook of child psychology and developmental science: Ecological settings and processes* (pp. 175–222). John Wiley & Sons, Inc.

Rubin, K., Fredstrom, B., & Bowker, J. (2008). Future directions in… friendship in childhood and early adolescence. *Social development (Oxford, England)*, *17*(4), 1085.

Rubin, K. H., Wojslawowicz, J. C., Rose-Krasnor, L., Booth-LaForce, C., & Burgess, K. B. (2006). The best friendships of shy/withdrawn children: Prevalence, stability, and relationship quality. *Journal of Abnormal Child Psychology*, *34*(2), 139–153.

Savin-Williams, R. C., & Berndt, T. J. (1990). Friendship and peer relations. In S. S. Feldman, & G. R. Elliott (Eds.), *At the threshold: The developing adolescent* (pp. 277–307). Cambridge, MA: Harvard University Press.

Schacter, H. L., Lessard, L. M., Kiperman, S., Bakth, F., Ehrhardt, A., & Uganski, J. (2021). Can friendships protect against the health consequences of peer victimization in adolescence? A systematic review. *School Mental Health*, 1–24.

Schneider, B. H., Fonzi, A., Tani, F., & Tomada, G. (1997). A cross-cultural exploration of the stability of children's friendships and the predictors of their continuation. *Social Development*, *6*(3), 322–339. https://doi.org/10.1111/j.1467-9507.1997.tb00109.x.

Schwartz-Mette, R. A., Shankman, J., Dueweke, A. R., Borowski, S., & Rose, A. J. (2020). Relations of friendship experiences with depressive symptoms and loneliness in childhood and adolescence: A meta-analytic review. *Psychological Bulletin*, *146*(8), 664.

Selman, R. L. (1980). *The growth of interpersonal understanding: Developmental and clinical analyses*. San Diego: Academic Press.

Shell, M. D., Gazelle, H., & Faldowski, R. A. (2014). Anxious solitude and the middle school transition: A diathesis × stress model of peer exclusion and victimization trajectories. *Developmental Psychology*, *50*(5), 1569–1583. https://doi.org/10.1037/a0035528.

Shin, H. (2017). Friendship dynamics of adolescent aggression, prosocial behavior, and social status: The moderating role of gender. *Journal of Youth and Adolescence*, *46*(11), 2305–2320.

Sijtsema, J. J., & Lindenberg, S. M. (2018). Peer influence in the development of adolescent antisocial behavior: Advances from dynamic social network studies. *Developmental Review*, *50*, 140–154.

Sprecher, S. (2002). Sexual satisfaction in premarital relationships: Associations with satisfaction, love, commitment, and stability. *Journal of Sex Research*, *39*(3), 190–196.

Steinhoff, A., & Keller, M. (2020). Pathways from childhood sociomoral sensitivity in friendship, insecurity, and peer rejection to adult friendship quality. *Child Development*, *91*(5), e1012–e1029.

Sullivan, H. S. (1953). *The interpersonal theory of psychiatry*. New York: Norton.

Troop-Gordon, W., & Ladd, G. W. (2005). Trajectories of peer victimization and perceptions of the self and schoolmates: Precursors to internalizing and externalizing problems. *Child Development*, *76*(5), 1072–1091.

Urberg, K. A. (1992). Locus of peer influence: Social crowd and best friend. *Journal of Youth and Adolescence*, *21*(4), 439–450.

Urberg, K. A., Değirmencioğlu, S. M., & Pilgrim, C. (1997). Close friend and group influence on adolescent cigarette smoking and alcohol use. *Developmental Psychology*, *33*(5), 834.

Van Zalk, N., Van Zalk, M., Kerr, M., & Stattin, H. (2011). Social anxiety as a basis for friendship selection and socialization in adolescents' social networks. *Journal of Personality*, *79*(3), 499–526.

Vaughn, B. E., Colvin, T. N., Azria, M. R., Caya, L., & Krzysik, L. (2001). Dyadic analyses of friendship in a sample of preschool-age children attending head start: Correspondence between measures and implications for social competence. *Child Development*, *72*(3), 862–878.

Way, N., & Greene, M. L. (2006). Trajectories of perceived friendship quality during adolescence: The patterns and contextual predictors. *Journal of Research on Adolescence*, *16*(2), 293–320.

Wojslawowicz Bowker, J. C., Rubin, K. H., Burgess, K. B., Booth-LaForce, C., & Rose-Krasnor, L. (2006). Behavioral characteristics associated with stable and fluid best friendship patterns in middle childhood. *Merrill-Palmer Quarterly*, *52*, 4.

You, J., Lin, M. P., Fu, K., & Leung, F. (2013). The best friend and friendship group influence on adolescent nonsuicidal self-injury. *Journal of Abnormal Child Psychology*, *41*(6), 993–1004.

CHAPTER TEN

The development of metacognitive knowledge from childhood to young adulthood: Major trends and educational implications

Wolfgang Schneider*, **Catharina Tibken, and Tobias Richter**
Department of Psychology, University of Würzburg, Würzburg, Germany
*Corresponding author: e-mail address: schneider@psychologie.uni-wuerzburg.de

Contents

Abstract

In this chapter, major trends in the development of metacognitive competences and its relationship to various aspects of cognitive/academic performance are described, with a focus on metamemory and reading comprehension. In a first step, classic and more recent theoretical models of metacognitive competences are presented that elaborate on the development of the declarative and procedural components of metacognition. The declarative component focuses on explicit and verbalizable knowledge, whereas the procedural component deals with monitoring and self-regulation processes. Common measures of both components are presented next, followed by a description of major developmental trends, as indicated by cross-sectional and longitudinal studies

Advances in Child Development and Behavior, Volume 63
ISSN 0065-2407
https://doi.org/10.1016/bs.acdb.2022.04.006

on developmental differences and changes in declarative knowledge, memory and comprehension monitoring, and self-regulation. Furthermore, research findings illustrating the relationship between metacognitive competences and cognitive performance in childhood and adolescence are presented for the domains of memory and reading comprehension, respectively. The final section focuses on educational implications of research on metacognition, underlining the importance of teacher behavior in the classroom and special instruction programs for students' acquisition of metacognitive competences.

During the last five decades, numerous publications have focused on the development of memory, mostly in children and adolescents (for reviews, see Bauer & Fivush, 2014; Schneider, 2015; Schneider & Ornstein, 2019). Overall, developmental changes in memory capacity, memory strategies, domain-specific knowledge, and metacognitive competences were assumed to cause increases in memory performance across childhood and adolescence. In a first series of studies carried out from the 1970s on, the contribution of memory strategy development to improvements in memory performance was assessed and evaluated for different age groups (see Roebers, 2014; Schneider & Bjorklund, 1998). Subsequently, many researchers in this field investigated the relative impact of metacognitive competences on memory performance more thoroughly. In addition, research on metacognitive competences was extended to other performance domains such as reading comprehension or mathematical competences.

In this chapter, we focus on two important performance domains, in which metacognitive competences play a pivotal role: memory and reading comprehension. In the following, we will present conceptualizations of the theoretical construct of metacognition and research findings on the assessment and development of metacognitive competences. At the end of this chapter, the role of metacognitive competences in the development of memory and academic achievement (focusing on reading comprehension) will be discussed.

1. Conceptualizations and models of metacognitive competences

Research on metacognitive development was initiated in the early 1970s by Ann Brown, John Flavell and their colleagues in the domain of memory research (for reviews, see Brown, Bransford, Ferrara, & Campione, 1983; Flavell, Miller, & Miller, 2002; Schneider & Pressley, 1997). Flavell (1971) first introduced the term *metamemory* to refer to knowledge about memory processes and contents. In their taxonomy of metamemory, Flavell and Wellman (1977) distinguished between two main categories, "sensitivity" and "variables." The "sensitivity" category referred to mostly implicit, unconscious behavioral knowledge of when memory is necessary, and thus was very close to subsequent conceptualizations of *procedural metacognitive competences*. The "variables" category referred to explicit, conscious, and factual knowledge about the importance of person, task, and strategy variables for memory performance. This is also known as *declarative metacognitive knowledge*. Declarative metacognitive knowledge about *person variables* includes knowledge about how, when, and why one remembers or forgets. Declarative metacognitive knowledge about *task variables* comprises knowledge about task influences on memory performance, for instance, knowledge that shorter item lists are easier to remember than long lists. Finally, declarative metacognitive knowledge about *strategy variables* refers to knowledge about advantages and possible problems of memory strategies. The same classification of declarative metacognitive competences can also be applied to other domains such as reading comprehension. In the domain of reading comprehension, for example, a person variable refers to knowledge about individual strengths and weaknesses in reading comprehension, a task variable refers to metacognitive knowledge about text difficulty, and a strategy variable refers to knowledge about reading strategies.

This taxonomy of metacognition was not intended to be exhaustive. Since the late 1970s, a number of other researchers have contributed to the development of theories about metacognitive competences, mostly in the field of metamemory (for reviews, see Joyner & Kurtz-Costes, 1997; Schneider, 2015; Schneider & Löffler, 2016). For instance, Paris and colleagues (e.g., Paris & Oka, 1986) introduced a component called "conditional metacognitive knowledge" that focused on children's ability to justify or explain their decision concerning memory activities. Whereas the declarative metamemory component introduced by Flavell and Wellman focused on "knowing that," conditional metamemory referred to "knowing why" (see also Schraw & Moshman, 1995).

The procedural component of metamemory, that is, children's ability to *monitor* and *regulate* their memory behavior ("knowing how") was first analyzed by Ann Brown and colleagues (e.g., Brown et al., 1983). An impression that was derived from some of Flavell's early research was that a lot of metacognitive development was complete by age 8 or 9 (e.g., Kreutzer, Leonard, & Flavell, 1975), and that this was true for both declarative and procedural metacognitive competences. One reason Ann Brown and her colleagues were motivated to re-conceptualize metamemory was to counteract this impression. They focused on the competent information processor, one possessing an efficient "executive" that regulated cognitive behaviors. In their view, this regulatory component is responsible for selecting and implementing strategies, monitoring their usefulness, and modifying them when necessary. The assumption was that children do not monitor and regulate their performance as well as adolescents and adults do. Overall, Brown et al. (1983) adopted the perspective that memory monitoring and regulation processes play a large role in complex cognitive tasks such as comprehending and memorizing text materials. They also argued that declarative and procedural components of metamemory are fundamentally different in nature. Whereas the declarative knowledge component is primarily verbalizable, stable, and late-developing, the procedural component is not necessarily verbalizable, rather unstable, relatively age-independent, and is determined by the specific task or situation.

In the following years, metacognition was more broadly defined as any knowledge or cognitive activity that takes as its cognitive object, or that regulates, any aspect of any cognitive activity (Flavell, 1979; Flavell et al., 2002). Obviously, this conceptualization refers to people's knowledge of their own information processing skills, as well as knowledge about the nature of cognitive tasks, and about strategies for coping with such tasks. Moreover, it also includes executive skills related to monitoring and regulation of one's own cognitive activities.

Since the beginnings of metacognition research in the context of memory development, other research domains have adapted the concept. Text comprehension research is one of these domains. As children and adolescents grow older, they increasingly read texts for comprehension and learning. Thus, reading comprehension, especially comprehension of expository texts, is an important prerequisite for academic success (Savolainen, Ahonen, Aro, Tolvanen, & Holopainen, 2008). While reading texts, readers use the text information and their prior knowledge to successively build a coherent mental representation of the text content, called a mental model or situation model (Zwaan & Radvansky, 1998). Building a mental model during comprehension is to a large degree an automated and passive process (e.g., Myers & O'Brien, 1998; Richter, 2015). However, depending on their reading goal, readers apply different standards of coherence (van den Broek & Helder, 2017). When readers have high standards of coherence (e.g., when learning from a text), they do not only rely on passive processes but also engage in active metacognitive processes to improve their comprehension.

Comparable to the domain of metamemory (Brown et al., 1983), these procedural metacognitive processes can be distinguished into processes of (comprehension) monitoring and (comprehension) regulation. Comprehension monitoring includes the detection of coherence breaks, such as logical inconsistencies within the text. Comprehension regulation refers to the (more or less conscious) control of reading processes in order to resolve perceived comprehension difficulties, for example by slowing-down the reading process, by reading difficult text passages repeatedly, or by looking up unknown words. A more general theoretical framework was established by Pressley, Borkowski, and their colleagues (e.g., Pressley, Borkowski, & Schneider, 1989), who proposed an elaborate model of metacognition, the Good-Information-Processing-Model, that not only considered aspects of procedural and declarative metacognitive competences but also linked these concepts to other features of successful information processing. According to this model, sophisticated metacognition is closely related to the learner's strategy use, motivational orientation, general knowledge about the world, and automated use of efficient learning procedures. All of these components are assumed to interact. For instance, specific strategy knowledge (declarative aspect) influences the adequate application of metacognitive strategies (procedural aspect). As the strategies are carried out, they are monitored and evaluated, which—in turn—leads to expansion and refinement of specific strategy knowledge.

Other conceptualizations of metacognition added components such as self-regulation skills and implicit metacognitive experiences (e.g., Efklides, 2008, 2019; Kuhn, 2000). In her model, Efklides emphasized the importance of specific learning-related metacognitive experiences (comparable to metacognitive experience in the Flavell's model), which should play a crucial role in the development of metacognitive skills. Accordingly, learning outcomes are influenced not only by conscious and explicit information processing skills but also by implicit metacognitive experiences that arise from a person's affective and motivational reaction to a given task. Similarly, Kuhn (2000) assumed that the acquisition of both declarative and procedural metacognitive competences is accompanied by predominantly unconscious associative processes during learning. In the beginning, the emergence of metacognitive competences is implicit, gradually becomes explicit, and can eventually be accessed consciously.

Whereas the concept of metacognition was first developed in the context of developmental (memory) research, it is now widely used in different areas of psychology, including motivation research, clinical and educational psychology. More recent developments also include cognitive neuroscience models of metacognition (e.g., Fleming & Frith, 2014; Fleur, Bredeweg, & van den Bos, 2021). Its popularity is mainly due to the fact that metacognition is crucial for concepts of everyday reasoning and those assessing scientific thinking as well as social interactions.

2. Assessment of metacognitive competences

A variety of measures have been used to capture children's metacognitive competences. Measures assessing *declarative metacognitive competences* are taken without concurrent assessment of cognitive performance, whereas measures of procedural metacognitive competences are collected while working on a specific cognitive task (concurrent measures). Most measures of declarative, factual knowledge have utilized interviews or questionnaires that focus on general knowledge about person variables, task demands, and strategies. Whereas earlier instruments predominantly used in the memory domain suffered from methodological problems (see Cavanaugh & Perlmutter, 1982), more recent interviews and questionnaires showed better psychometric properties such as sufficient reliability and validity (see Best & Folds-Bennett, 2021; Schneider, 2010). For instance, it was shown that nonverbal techniques helped in assessing young children's declarative knowledge. When the task required young children to distinguish effective from poor memory strategies while watching a model executing the strategies on video, many children were able to master this task. In another successful procedure ("peer-tutoring task"), older children were asked to tutor a younger child about how to do a certain memory task in order to maximize learning. Peer tutoring is likely more motivating to young children than interviews, and they tend to be more explicit when answering a question of an older child as compared with that of an adult (who already seems to know everything), which leads to more reliable measures of declarative metacognitive knowledge.

Overall, these alternative methods alleviated some of the problems usually related to the use of questionnaire measures, especially in young children. However, these measures still created difficulties when applied to older children and adolescents. There is a risk that social desirability factors reduce the validity of outcomes on declarative metacognitive competences for this target group. Accordingly, more sophisticated measures of declarative metacognitive competences have been used with older children and adolescents. For instance, in the domain of reading comprehension, Schlagmüller and Schneider (2007) came up with a standardized measure of metacognitive knowledge that was based on a revised test instrument developed for the International Program for Student Assessment (PISA 2000; see Artelt, Schiefele, & Schneider, 2001) and was later included in the PISA 2009 assessment (see below). This instrument taps adolescents' knowledge of reading strategies which are particularly relevant for reading comprehension and thus also for learning from texts. For each of six scenarios, students evaluate the quality and usefulness of five different strategies available for reaching the intended learning goal.

Procedural metacognitive competences have been assessed by measures that are characterized by the presence of a simultaneous cognitive activity such as memorizing a word list or comprehending a text. We will first present typical measures used in memory research and then describe the assessment of procedural metacognitive competences in the domain of reading comprehension.

The most intensely studied aspect of procedural metacognitive competence in the domain of memory research is *memory monitoring*, that is, evaluating how well one is progressing in a memory task. An often-used approach is to ask children and adolescents to judge their (future or past) memory performance either shortly before, during, or after working on a memory task (for overviews, see Brown et al., 1983; Flavell et al., 2002; Roebers, 2002; Schneider & Lockl, 2008).

Ease-of-learning (EOL) judgments occur in advance of the learning process, are largely inferential, and refer to items that have not yet been learned (Nelson & Narens, 1994). Participants have to judge how easy it will be to learn a given set of items. The corresponding memory paradigm is performance prediction. In comparison, *judgments-of-learning* (*JOLs)* occur, during or soon after, the acquisition of memory materials and are predictions about future test performance on recently studied (and probably still recallable) items. Typically, paired-associate learning tasks are used in this context. After completion of a learning trial, participants are shown the stimuli of a given pair and have to indicate how confident they are whether they will remember the correct item response either immediately or a certain time later. A number of developmental studies also explored children's *feeling-of-knowing (FOK) judgments* (e.g., Lockl & Schneider, 2002). These judgments occur either during or after a learning procedure and are judgments about whether a currently unrecallable item will be remembered at a subsequent retention test (when a selection of possible answers is presented). Finally, *confidence judgments (CJs)* are used to assess monitoring at retrieval. After having provided an answer to a memory question, participants are asked to indicate how sure they are that the answer is correct.

Whereas memory monitoring involves knowing where you are with regard to your goal of understanding and memorizing task materials, *memory regulation* includes planning, directing, and evaluating one's mnemonic activities (Flavell et al., 2002). Some developmental studies in the domain of memory research addressed aspects of children's control and self-regulation processes such as termination of study (recall readiness) and allocation of study time (see the reviews by Schneider, 2015; Schneider & Löffler, 2016). *Recall readiness* assessments are made after learning materials have been studied at least once. Typically, participants are asked to continue studying until their memory of the materials to be learned is perfect. Another example of regulation skills concerns the *allocation of study time*. This research observes how learners deploy their attention and effort when studying lists of items. For instance, developmental studies on the allocation of study time examined whether school children and adults tend to spend more time on less well-learned material. After a first free recall trial, participants had to distinguish between recalled and non-recalled items (monitoring component) and were then asked to select half of the items for additional study (regulation component). One problem with the paradigm of the allocation of study time is that it may not only tap metacognitive processes but may also be influenced by motivational variables (see Schneider & Lockl, 2002).

A number of methodological concerns have been raised about measures of procedural metamemory. For instance, problems with measuring *calibration*, that is, the accuracy of metacognitive judgments, has been repeatedly discussed (e.g., Boekaerts & Rozendaal, 2010; Lingel, Lenhart, & Schneider, 2019; Schraw, Kuch, Gutierrez, & Richmond, 2014). However, although these evaluations indicate that there is not a perfect index of metamemory, many of the measurement problems that metamemory researchers confront are similar to measurement problems in other areas of psychology.

In the domain of reading comprehension, like in metamemory research, calibration measures have been used to capture procedural metacognitive competences (e.g., de Bruin, Thiede, Camp, & Redford, 2011). Inconsistency tasks are another frequently used measure to assess procedural metacognitive competences during and after reading (e.g., Baker & Anderson, 1982; Helder et al., 2016; Tibken, Richter, Wannagat, et al., 2022; Tibken, Richter, von der Linden, Schmiedeler, & Schneider, 2022). Inconsistencies are logical contradictions within a text or contradictions of text information with world knowledge. These inconsistencies disrupt the reader's flow of comprehension (Baker, 1989). A discrepancy between incoming information during reading and the mental representation of the text content up to that point thus triggers comprehension monitoring, provided that readers have high standards of coherence (van den Broek, Risden, & Husebye-Hartman, 1995), for example, because they read for study purposes. The results of this comprehension monitoring may in turn lead to effortful and time-consuming processes of comprehension regulation to resolve the inconsistencies, such as repeated reading of inconsistent sentences (Zabrucky & Ratner, 1986).

The inconsistency task can be applied to capture comprehension monitoring and—to some part—regulation offline (after reading) and online (during reading). To measure comprehension monitoring offline, participants are required to explicitly state whether an inconsistency had occurred in a previously read text, for example, by pointing out the inconsistency. To measure comprehension monitoring online, reading times for inconsistent sentences are compared to reading times for consistent sentences (e.g., Helder et al., 2016; Zabrucky & Ratner, 1986) or eye-movements are used to capture disfluencies caused by inconsistent information (van der Schoot, Reijntjes, & van Lieshout, 2012). Longer reading times for inconsistent compared to consistent sentences may be interpreted as an (implicit) indicator of comprehension monitoring, given that inconsistencies may disrupt the reading process and readers' attempts to resolve the inconsistency take time (e.g., Grabe, Antes, Kahn, & Kristjanson, 1991; O'Brien, Rizzella, Albrecht, & Halleran, 1998). Performance in the offline and in the online measure is often related, although readers are not always able to report on inconsistencies that slowed down their reading. However, in some studies, longer reading times for inconsistent compared to consistent target sentences occurred predominantly when the participants explicitly stated that they had noticed an inconsistency (Helder et al., 2016; Münchow, Richter, von der Mühlen, & Schmid, 2019; Zabrucky & Ratner, 1986).

A methodological limitation of the inconsistency task is that it is impossible to dissociate comprehension monitoring and regulation. The offline measure reflects the result of prior comprehension monitoring as participants are asked report previously detected inconsistencies after reading, whereas the online measure provides little information about the extent to which it captures aspects of monitoring and regulation. Longer reading times when reading inconsistent (compared to consistent) sentences might reflect comprehension monitoring, which leads to higher cognitive demands and consequently to a slow-down during reading, but it might also partly reflect an attempt to improve and thus regulate comprehension by rereading sentences or text passages. Despite these limitations, the inconsistency task has proven to be a suitable measure for examining procedural metacognitive competences in the domain of reading comprehension.

3. Development of metacognition

3.1 Precursors of metacognitive competences

Since the early 1980s, there has been study of preschoolers' metacognition motivated by Perner's (1991) and Wellman's (1990) conceptualizations of children's *theory of mind*. These studies emphasize that by the ages of 3–4 children have acquired important classes of knowledge about the inner mental world. From this age on, children develop a rudimentary understanding of mental verbs such as "thinking" or "remembering" and can separate mental processes from external behaviors associated with them. Children then gradually learn to recognize that the mental world can be differentiated into processes such as remembering, knowing, and guessing (that is, they acquire knowledge about distinct mental processes). Three- to four-year-olds are not generally capable of differentiating these processes, whereas older preschoolers already make distinctions that are very similar to those of adults.

The relationship between the development of language, theory of mind development, and (declarative) metamemory development has been systematically explored in longitudinal studies. For instance, Astington and Jenkins (1999) found that language competence predicted theory of mind development, but not the reverse. Longitudinal relationships among metacognitive vocabulary (knowledge of mental verbs), theory of mind, and subsequent metamemory were analyzed by Lockl and Schneider (2007). It was shown that both early theory of mind competence and metacognitive vocabulary affected subsequent metamemory. Findings also indicated that metacognitive vocabulary, theory of mind, and general metamemory improved considerably over the preschool years. Moreover, both early theory of mind and metacognitive vocabulary substantially predicted metamemory at the end of the preschool period even when individual differences in nonverbal intelligence and general vocabulary were taken into account. Overall, however, mean performance in the metacognitive vocabulary and metamemory assessments were far from the ceiling, indicating that metacognitive knowledge is not particularly rich before children enter school.

3.2 Children's declarative knowledge about memory

The first comprehensive questionnaire to assess children's declarative metamemory was created in a classic study by Kreutzer et al. (1975). Children in kindergarten, Grades 1, 3, and 5, were asked about person, task, and strategy variables. For example, children were asked if they ever forgot things, if it was easier to recall the gist of a story than to recall it verbatim, and what they could do to find a jacket they had lost while at school. Overall, the results of this study and related assessments (e.g., Myers & Paris, 1978; Schneider, Kron, Hünnerkopf, & Krajewski, 2004) indicated substantial improvements on most of the variables as a function of age. The fifth graders in the Kreutzer et al. study displayed a more sophisticated understanding of the situations and actions that predict successful memory than did the third graders and kindergarten children.

Sodian and Schneider (1999) analyzed the development of declarative metacognitive knowledge about memory strategies using data from the Munich Longitudinal Study (LOGIC). At 4 years of age, participants were given a short interview that addressed the utility of various memory strategies. This interview was repeated when children were 6, 8, 10, and 12 years of age. As can be seen from Fig. 1, most 6-year-olds chose sorting by color as

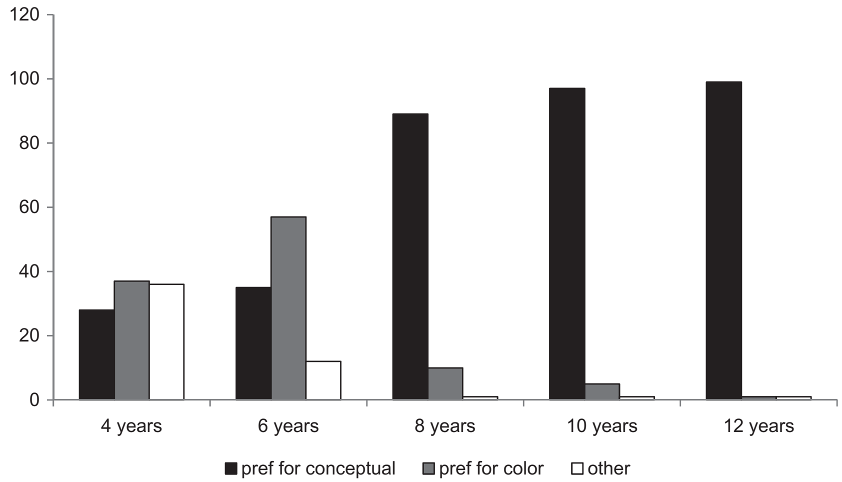


Fig. 1 The development of task-related metamemory. *Adapted from Sodian, B., & Schneider, W. (1999). Memory strategy development—Gradual increase, sudden insight or roller coaster? In F. E. Weinert & W. Schneider, W. (Eds.),* Individual development from 3 to 12: Findings from the Munich Longitudinal Study *(pp. 61–77). Cambridge University Press.*

the best memorization strategy, whereas sorting by semantic category was judged as the best strategy from age 8 on. There was also some evidence that individual differences in declarative metacognitive knowledge about memory strategies were related to further stability in strategy use.

Although clear age trends have been demonstrated in several studies, this does not mean that young children do not possess any adequate knowledge about memory. Even the kindergarten children in the Kreutzer et al. (1975) study knew that remembering many items is more difficult than remembering just a few, and the majority of these children also knew that using external devices (e.g., writing telephone numbers down) helps in remembering. Young children do have a basic understanding of memory, but declarative metacognitive knowledge about the importance of task characteristics and memory strategies develops more rapidly once children enter school. Knowledge about the usefulness of memory strategies was tapped in several studies that focused on organizational strategies (see reviews by Hasselhorn, 1990; Joyner & Kurtz-Costes, 1997; Schneider, 2015). Preferences for the most appropriate (sorting and clustering) strategies were not found before the ages of 8 or 10, and reasonable justifications for such preferences were not always provided.

Similar age trends were observed when children had to consider several dimensions of memory performance simultaneously (interactive memory knowledge). In a classic study, Wellman (1978) presented memory problems to 5- and 10-year-olds. Each problem consisted of ranking three picture cards, each of which contained a memorizing scenario. All of the children solved the simple problems tapping a single task variable such as the impact of number of items on memory performance, whereas substantial developmental differences were found for the complex memory problems varying two aspects (e.g., number of items and type of strategy). Only a very small proportion of the younger age group were able to judge the complex memory problems appropriately. The available data indicate that interactive memory knowledge develops very slowly, a process continuing well into adolescence (see Schneider & Pressley, 1997).

Taken together, the empirical evidence illustrates important changes in declarative metamemory over time. Using sensitive methods that minimize demands on the children, it has been possible to demonstrate some rudimentary declarative knowledge about memory functioning in preschoolers. Declarative metacognitive knowledge about memory develops rapidly during the course of elementary school and is already impressive by 11 or 12 years of age (cf. Roebers, 2014; Schneider & Löffler, 2016).

Nonetheless, declarative metacognitive competences are not complete by the end of childhood. It seems important to note that even though metacognitive knowledge increases substantially between young childhood and young adulthood, there is also evidence that many adolescents (including college students) demonstrate little knowledge of powerful and important strategies when the task is to read, comprehend, and memorize complex text materials (cf. Brown et al., 1983; Garner, 1987; Pressley & Afflerbach, 1995). Edossa, Neuenhaus, Artelt, Lingel, and Schneider (2018) showed that from Grade 5 to 8 there are still substantial gains in declarative metacognitive knowledge about reading strategies.

3.3 Development of procedural metacognitive competences

Memory monitoring. As noted above, most developmental studies on procedural metamemory focused on memory monitoring such as EOL, JOL, FOK and CJ, and also addressed their relation to regulation skills. Compared with the substantial age trends demonstrated for declarative metacognitive knowledge, the developmental pattern regarding procedural metamemory is less clear. What are the major developmental trends? In short, the findings suggest that even young children possess monitoring skills, and that developmental trends are not entirely clear, varying as a function of the paradigm under study. Whereas young kindergarten children tend to overestimate their performance when EOL judgments are considered, performance can be already accurate in young elementary school children. There is evidence that young children's overestimations of future performance are not due to metacognitive deficiencies but are at least partially caused by children's wishful thinking and their belief that effort has a powerful effect on performance (see Schneider & Lockl, 2008). When children's post-dictions were assessed in children ranging between 7 and 10 years of age, rather accurate judgments were found even for the younger age groups. In most studies, only subtle improvements over the elementary school years were found (for reviews, see Joyner & Kurtz-Costes, 1997; Roebers, 2002; Schneider & Löffler, 2016).

Given that only a few developmental studies focused on judgments of learning (JOLs) occurring during or soon after the acquisition of memory materials, the evidence regarding developmental trends is scarce. Overall, findings support the assumption that children's ability to judge their own memory performance after study of test materials seems to increase over the elementary school years. However, even young elementary school children are able to monitor their performance quite accurately when judgments are not given immediately after study but are somewhat delayed (Lipowski, Merriman, & Dunlosky, 2013; Roebers, von der Linden, Howie, & Schneider, 2007; Schneider, Visé, Lockl, & Nelson, 2000). Immediate JOLs are typically inaccurate, regardless of age, and in most cases represent overestimations of actual performance.

A number of studies have explored children's feeling-of-knowing (FOK) judgments and accuracy. Overall, most of the early studies on FOK judgments suggested that FOK accuracy improves continuously across childhood and adolescence. However, more recent investigations indicate that the pattern of developmental trends may be different, with the progression in children's JOLs and FOK being rather slow initially but accelerating when children enter school. Overall, the empirical evidence suggests an asymptotic function, with monitoring related to encoding processes already being rather accurate by the age of 8 years (see Roebers, 2014; Schneider & Löffler, 2016).

As to retrieval monitoring (confidence judgments), many relevant studies were carried out in naturalistic contexts, particularly those exploring children's eyewitness memory. After having provided an answer, participants were asked to indicate how sure they were that the answer was correct. CJs are thought to reflect a substantive sense of certainty that arises from the strength of memory that is being retrieved. The latter has been interpreted as an indicator of memory accuracy (Ghetti, Lyons, Lazzarin, & Cornoldi, 2008). In most studies on retrieval monitoring, developmental trends were found in that the overall level of confidence decreased with age, indicating that the degree of overconfidence typically found for children is reduced in adults. Moreover, the differentiation between correct and incorrect answers gets more pronounced with increasing age. That is, CJs are generally more accurate in older school children and adolescents because they feel more uncertain than younger children in the case of incorrect responses (Roebers, 2002; von der Linden & Roebers, 2006). Correlations between monitoring of encoding (FOK judgments, JOLs) and retrieval monitoring (CJs) assessed in the study by Roebers et al. (2007) were only moderate for younger school children (8-year-olds) but increased with age. Apparently, the various memory monitoring competencies described above get more consistent and reliable in late childhood and adolescence.

Memory regulation. Studies exploring developmental trends in *memory regulation* (as well as the interaction between monitoring and regulation) revealed a more clear-cut developmental pattern, indicating considerable increases from middle childhood to adolescence. Spontaneous and effective use of regulation skills occurred only in highly constrained situations during the grade-school years and continued to develop well into adolescence. Comparisons of younger and older children in "study-time apportionment" tasks indicate that it is the interplay between monitoring and regulatory activities that develops with age. That is, when the task is to learn item pairs of varying difficulty, both younger and older children show adequate monitoring skills in that they are well able to distinguish difficult from easy item pairs. However, only the older children allocate study time differentially, spending more time on the difficult than on the easier items. In comparison, younger children typically spend about the same amount of time on easy pairs as they spend on difficult pairs. The available studies thus indicate an age-related improvement in the efficient allocation of study time. Developmental differences were not so much observed in the metacognitive knowledge itself but in its efficient application to regulation strategies (Schneider & Lockl, 2008).

Relations between memory monitoring and memory regulation. A very common assumption in metacognitive research is that monitoring processes influence regulation and learning behavior (e.g., Son & Metcalfe, 2000), In fact, numerous developmental studies supported this "monitoring-affects-control" (MC) model. According to this model, learners first monitor item difficulty before actually investing study time. The outcome of the monitoring processes then serves as a basis for the allocation of study time. However, more recent research has challenged this position, suggesting that regulatory processes may also influence metacognitive monitoring, which supports the assumption of a "control-affects-monitoring" (CM) model (see Koriat, 2008). For instance, Koriat, Ackerman, Lockl, and Schneider (2009) proposed that study time is actually used by learners as a cue for encoding fluency under what they called the *memorizing effort heuristic.* Accordingly, easily studied items are considered to be more likely remembered than items that require more effort to study. Thus, metacognitive judgments are basically data-driven: study time duration is taken retrospectively as a cue for the feeling of mastery. Greater effort (longer study time) is associated with lower JOLs, suggesting that the cause-and-effect relation is actually from control to monitoring. The results of Koriat et al. (2009) not only confirmed this assumption for older school children but also showed evidence for the "monitoring-affects-control" hypothesis: monitoring affects study time when regulation is goal-driven, that is, when the intention to reach a specific goal is motivated by incentives or rewards. Obviously, the relationships between monitoring and control can be more complex than originally assumed. More recent developmental studies on this issue using samples of children and adults explored whether (dependent on the instruction) "data-driven" and "goal-driven" regulation can be observed within the same paired-associate learning task (see Koriat, Ackerman, Adiv, Lockl, & Schneider, 2014; von der Linden, Löffler, & Schneider, 2017). Both studies confirmed the assumption that this is possible for adolescents and adults, but not for elementary school children. Thus, the ability to switch between the "monitoring-affects-control" and "control-affects-monitoring" models develops rather late.

Metacomprehension. To examine the development of procedural metacognitive competences in the domain of reading comprehension, studies have focused on age differences in performance on the inconsistency task as an indicator of comprehension monitoring. Hacker (1997) found an increase regarding the number of correctly detected inconsistencies in students from Grade 7 to Grade 11 (explicit aspect of comprehension monitoring). Accordingly, Helder et al. (2016) showed that students in Grade 5 detected significantly more inconsistencies than students in Grade 3.

In contrast, Helder et al. (2016) found no difference in the slow-down of reading times for inconsistent compared to consistent sentences (implicit aspect of comprehension monitoring) between students in Grade 3 and 5. A study by Zabrucky and Ratner (1986) with students in Grades 3 and 6 indicated a comparable pattern of results, with age-related differences for explicit detection of inconsistencies but not for the implicit measure of reading time differences. In contrast to the results by Helder et al. (2016) and Zabrucky and Ratner (1986), Zargar, Adams, and Connor (2019) found a greater slow-down in students in Grade 5 than Grade 4 in reading times for inconsistent compared to consistent sentences. Students in Grade 3 showed no inconsistency effect in their reading times at all. These findings indicate that young elementary school children fail to detect inconsistencies in texts, even when comprehension monitoring is assessed implicitly. In a study by Wassenburg, Bos, de Koning, and van der Schoot (2015), children in Grade 4 also showed greater differences in reading times between inconsistent and consistent sentences than children in Grade 3.

To conclude, studies on comprehension monitoring suggest that the ability to consciously perceive and report inconsistencies as an indicator of explicit comprehension monitoring considerably increases with age during childhood and adolescence. Findings on more implicit measures of comprehension monitoring (and in part regulation) that do not require participants to verbalize detected inconsistencies seem inconclusive. Differences between age groups might also depend on text characteristics and students' reading competences. For example, Connor et al. (2015) found that performance in the online aspects of comprehension monitoring improved during adolescence only in good readers. An eye-tracking study that also compared good and poor readers in Grade 5 and 6 found that poor readers had significantly more problems (shown in their reading times) when contradicting information was presented in passages that were widely separated within the text. In texts with contradicting information in successive sentences, there was no difference in the implicit measure of comprehension monitoring between good and poor readers (van der Schoot et al., 2012). Additionally, comprehension monitoring is easier for children in narrative than in expository texts (Currie et al., 2020).

With regard to comprehension regulation, Zabrucky & Ratner, 1986 found that students in Grade 6 looked back to previously read inconsistencies more frequently than students in Grade 3. However, only few studies investigated comprehension regulation in children. Findings on adults suggest that regulatory activity is more likely to occur when a text contradicts world knowledge than when it is self-contradictory (van Moort, Koornneef, & van den Broek, 2021).

Thus, comparable to findings on the development of procedural metacognitive competences in the domain of memory research, the development of comprehension monitoring and regulation during childhood is still not entirely clear, especially since the ability to monitor and regulate text comprehension seems to be affected by both individual reader and text characteristics.

4. Relations between metacognitive competences and cognitive performance

In this section, we discuss findings on the relations between metacognitive competences and cognitive performance. We start with studies on the relationship of metamemory and memory, followed by studies on the relationship of metacomprehension and reading comprehension.

4.1 Metamemory-memory relations

Early empirical research on metamemory was stimulated by the belief that young children do not spontaneously use memory strategies because they are not familiar with memory tasks and unable to judge the advantages of strategic behavior. Metamemory researchers believed that this situation should change after children enter school and are confronted with numerous memory tasks. From a developmental and educational perspective, the metamemory concept thus seems well-suited to explain children's "production deficiencies" on a broad variety of memory tasks. Experience with such tasks should improve children's metamemory, which in turn should have an impact on subsequent memory behavior (i.e., strategy use). The assumption was that although links between metamemory and memory may be weak in early childhood, they should become much stronger with increasing age.

In fact, most studies found that metamemory-strategy associations were generally weak in preschoolers and kindergarten children and increase with age (Joyner & Kurtz-Costes, 1997; Schneider & Bjorklund, 1998). Narrative reviews and meta-analyses have shown that there are moderate, non-trivial quantitative associations between metamemory and memory. For instance, Schneider and Pressley (1997) reported an overall correlation of about 0.41 based on about 60 publications and more than 7000 children and adolescents.

There are several possible explanations for the fact that the link between metacognition and cognitive performance is not always as strong as one might think, and, undoubtedly, one of the major mediators is motivation. Empirical research has shown that metamemory-memory performance relationships are particularly strong when participants are highly motivated to achieve a certain goal. Under these circumstances, all available resources will be activated to cope with task demands, including declarative and procedural metacognitive competences (cf. Borkowski & Mutukrishna, 1995).

Evidence from longitudinal and training studies. In addition to exploring concurrent associations between metamemory and use of cognitive memory strategies, there has been considerable interest in the question of developmental linkages between metamnemonic understanding and later strategy deployment (see Schneider, 2015; Waters & Kunnmann, 2010). As noted above, a common assumption has been that children would not make use of strategies for remembering until they have adequate levels of metamnemonic understanding. Testing this assumption, however, requires within-individual information over time so as to determine whether knowledge of the use of strategies precedes their later use. Meanwhile, several longitudinal studies provided support for the linkage between (earlier) metamemory and (later) strategy use. For instance, findings from a short-term longitudinal study (Schlagmüller & Schneider, 2002) and a longitudinal project designed to examine the impact of teachers' language during instruction on children's developing memory skills (e.g., Grammer, Purtell, Coffman, & Ornstein, 2011) indicated that increases in declarative metacognitive knowledge affect subsequent memory behavior.

Schlagmüller and Schneider (2002) carried out a "microgenetic" longitudinal on memory strategy development that assessed children's strategy use and metacognitive knowledge repeatedly within a short time period of several weeks. They reported that children who acquired an organizational strategy over the course of the study showed increases in declarative metamemory well ahead of actually showing the strategy. Additional evidence for such a metamemory—memory behavior link comes from the longitudinal study by Grammer et al., 2011 based on a sample of first-grade children who were followed into the beginning of second grade. Repeated assessments of children's organizational strategy use and declarative metamemory were made to examine the development of these skills and their interrelations over time. Results from latent growth curve models revealed that declarative metamemory at earlier time points was predictive of subsequent strategy use. Findings further suggested that the acquisition of declarative metamemory not only precedes organizational strategy use but also influences the amount of strategy use that children exhibit at subsequent time points.

From the 1980s on, several researchers have incorporated declarative metamemory information into training programs designed to enhance children's strategy use and memory performance. For instance, in a training program carried out by Ghatala, Levin, Pressley, and Goodwin (1986), elementary school children were presented with paired-associate learning tasks. Before studying these item lists, children were trained to (a) assess their performance with different types of strategies, (b) to attribute performance differences to the use of different strategies, and (c) to use this information to select the best strategy for a task. As a main result, it was shown that even children as young as 7–8 years of age can be taught to judge the relative efficacy of memory strategies. The Ghatala et al. (1986) study as well as most other training studies showed that providing metacognitive information about the value of being strategic increases the likelihood that children acquire a strategy and use it later (for reviews, see Pressley & Hilden, 2006; Schneider & Pressley, 1997).

Evidence from multivariate analyses. As noted by Schneider (2015), longitudinal work has convincingly shown that the links between strategy usage, metamemory, and recall become stronger with increasing age. Moreover, it has been repeatedly found that there are stronger relations between knowledge about a specific strategy and the deployment of that strategy than between metamemory and recall performance, primarily because strategy use directed by metamemory is only one of several determinants (e.g., IQ, working memory capacity, the nonstrategic knowledge base) of performance. Nonetheless, metamemory has been a particularly useful predictor of children's strategy use.

From the early 1980s on, comprehensive field studies using multivariate experimental designs have been used to examine the complex relationships among declarative metamemory, domain knowledge, strategy use, and memory performance, as well as their relationship with other important variables such as intelligence, working memory capacity, and motivation (e.g., DeMarie, Miller, Ferron, & Cunningham, 2004; Kron-Sperl, Schneider, & Hasselhorn, 2008; Schneider, Schlagmüller, & Visé, 1998). For instance, Schneider et al. (1998) used structural equation modeling to assess the relationships among verbal IQ, working memory capacity, declarative metamemory, the use of a semantic organizational strategy, and recall in a sort-recall task in third- and fourth graders (see Fig. 2).

As can be seen from Fig. 2, findings indicate that declarative metamemory was affected by both verbal IQ and working memory capacity. There was only a modest direct contribution of declarative metamemory to the prediction of recall, whereas the indirect link via strategy use was much stronger. As a consequence, individual differences in declarative metamemory explained a large proportion in the variance of the recall data. Similar findings were also reported by DeMarie et al. (2004), who illustrated

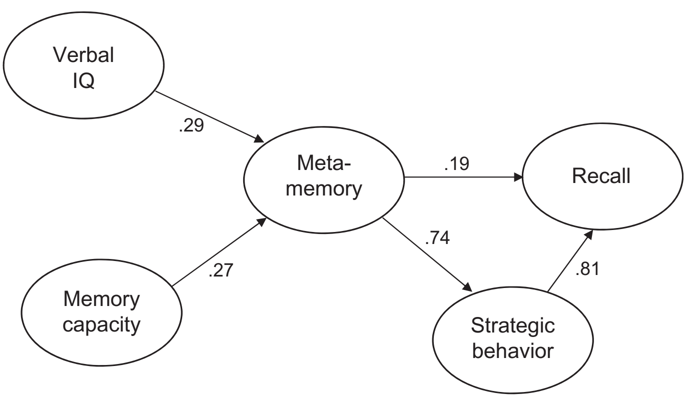


Fig. 2 Causal model depicting metamemory–memory relationships in a sort-recall task. *Adapted from Schneider, W., Schlagmüller, M., & Visé, M. (1998). The impact of metamemory and domain-specific knowledge on memory performance.* European Journal of Psychology of Education, 13*(1), 91–103. doi:10.1007/BF03172815.*

the importance of declarative metamemory for explaining individual differences in strategy use and memory performance in different age groups ranging from kindergarteners to fifth graders. DeMarie and colleagues also concluded that multiple strategy use must be considered, and that declarative metamemory plays a significant role in predicting the use of multiple strategies and strategy effectiveness. Whereas individual differences in memory capacity seemed to influence the use of strategies in younger children, individual differences in declarative metamemory more strongly affected strategy use in older children.

Overall, the findings of various multivariate analyses confirm that the causal relation between metacognitive knowledge and memory performance is complex. Metamemory sometimes has an indirect effect on recall, as when knowledge about categorization strategies leads to semantic grouping during the study period, which in turn produces better recall. Moreover, the influence seems to be bidirectional (see Flavell et al., 2002; Hasselhorn, 1995; Schneider & Pressley, 1997). That is, metacognition can influence cognitive behavior, which in turn leads to enhanced metacognition. Finally, many other relevant variables such as IQ, domain knowledge, memory capacity, and motivation significantly contribute to the explanation of differences in cognitive performance. Obviously, these empirical findings are in accord with the core assumptions of the Good-Information-Processing-Model (Pressley et al., 1989), which provides a detailed theoretical account of the complexity of these interactional processes (see also Section 1).

4.2 Relations between metacognitive competences and reading competence

*Declarative metacognitive competenc*es. Several large-scale studies explored the relationship between declarative metacognitive knowledge about reading strategies and reading comprehension. In a longitudinal study based on more than 900 German students, Schneider, Lingel, Artelt, and Neuenhaus (2017) assessed associations between metacognitive knowledge and academic achievement in reading, mathematics, and English (as a foreign language). Students were tested on six occasions from the beginning of Grade 5 until the end of Grade 9. Measures of declarative metacognitive knowledge similar to the one developed by Schlagmüller and Schneider (2007) were developed for each domain. The instrument developed for the domain of reading tapped students' knowledge of strategies relevant for comprehension as well as for recall of text information.

The findings of this study indicate that metacognitive knowledge develops substantially during the course of (German) secondary school. The growth processes in mathematics and reading assessed between Grades 5 and 9 were found to be negatively accelerated, indicating that more metacognitive knowledge was acquired at the beginning of secondary school than thereafter. In line with the existing literature, metacognitive knowledge also turned out to be an important predictor of achievement in secondary school students. A bi-directional relationship emerged between metacognitive and cognitive development in that the use of cognitive strategies improved the quality of metacognitive knowledge and improvement in metacognitive knowledge lead to a more sophisticated use of problem-solving strategies.

Artelt and Schneider (2015) used data of the OECD-PISA 2009 study to investigate the relationship between metacognitive knowledge, self-reported strategy use, and reading competence in 15-year-old students. An instrument tapping metacognitive knowledge about reading strategies similar to the one used by Schlagmüller and Schneider (2007) (see above) was included in the assessment, and the resulting indicator of metacognitive knowledge was related to students' reading competence scores. Moreover, the internationally comparative design allowed for an estimation of the cross-country generalizability of results. Data analyses were conducted for the reference group of the 34 OECD countries. Altogether, 298,454 15-year-old students participated, with subsample sizes for the 34 countries varying between 3664 (Iceland) and 38,250 (Mexico) students.

Correlational analyses revealed that although there was some variation among the coefficients between countries, the general picture that emerged was clear. As expected, moderate to high correlations between metacognitive knowledge and reading competence were found, ranging from 0.37 in Greece to 0.60 in Switzerland and Belgium, with an OECD mean correlation of 0.48. In comparison, the relationships between (self-reported) strategy use and reading competence were much lower. However, a moderating effect was found in most countries: For students with a rich knowledge base about when and where to apply strategies (conditional and relational strategy knowledge), the correlations between their self-reported use of control strategies and reading competence were higher than for students with a low knowledge base.

Procedural metacognitive competences. Most studies that examined relations between metacomprehension and reading competence cross-sectionally compared the performance of good and poor readers in, for example, the inconsistency task (e.g., Currie et al., 2020; Helder et al., 2016; Long & Chong, 2001; Oakhill, Hartt, & Samols, 2005; van der Schoot et al., 2012; Zabrucky & Ratner, 1992). These studies indicate that good readers perform better in explicit measures of comprehension monitoring than poor readers (e.g., Currie et al., 2020; Helder et al., 2016; Oakhill et al., 2005). Findings on associations of the implicit aspect of comprehension monitoring with reading competence are inconclusive. Some studies found no differences between good and poor readers in an implicit measure (e.g., Helder et al., 2016; Zabrucky & Ratner, 1989, 1992), whereas other studies reported better performance of good compared to poor readers (Ehrlich, Remond, & Tardieu, 1999). With regard to comprehension regulation, good readers reread difficult text passages more frequently than poor readers (Zargar et al., 2019).

A study that followed children from Grade 3 to 6 examined long-term relations between performance in an inconsistency task and reading competence (Oakhill & Cain, 2012). The results indicate cross-lagged effects between comprehension monitoring and reading competence in late childhood and early adolescence. Another longitudinal study with students in (initially) Grades 6 and 8 found a small significant effect of comprehension monitoring (detection of inconsistencies) on the development of reading competence within one school year, even if the effect of intelligence was controlled for (Tibken, Richter, von der Linden, et al., 2022). Comparable to the results with regard to metamemory and memory performance, the findings suggest that declarative metacognitive knowledge (about reading strategies) predicts procedural metacognitive competences (comprehension monitoring), which in turn affects reading competence. The effect of declarative metacognitive competences on procedural metacognitive competences was moderated by motivational characteristics (see Fig. 3). Thus, knowledge about reading strategies and a certain motivation

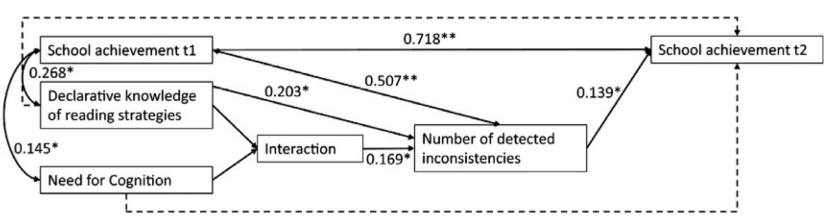


Fig. 3 Causal model depicting how declarative knowledge of reading strategies and need for cognition interactively predict school achievement via the number of detected inconsistencies on school achievement at t2 ($n = 129$ students in Grade 6/8, standardized path coefficients). *Adapted from Tibken, C., Richter, T., von der Linden, N., Schmiedeler, S., & Schneider, W. (2021). The role of metacognitive competences in the development of school achievement among gifted adolescents.* Child Development, 93*(1), 117–133. doi:10.1111/cdev.13640.*

to apply them when reading difficult text seems to lead to better comprehension monitoring (and probably regulation) and consequently to an improvement in reading competence.

5. Metacognition and education

During the last four decades, numerous studies have explored the development of metacognitive knowledge and the efficiency of strategy instruction approaches in school settings in the domain of memory research (for reviews, see Pressley & Hilden, 2006; Schneider, 2015). The basic assumption was that memory and metamemory development are not so much products of age but of education. Thus, several studies explored the importance of educational experiences for children's use of memory strategies and their metacognitive knowledge.

5.1 The role of teachers

Kron-Sperl et al. (2008) provide an example of the impact of educational experience on the development of memory and metacognitive knowledge in their longitudinal study. In this study, children were followed from kindergarten age to the end of elementary school. Kron-Sperl and colleagues repeatedly presented the children of their sample with a sort-recall memory task without giving any specific strategy cues. When performance of these children at the end of the longitudinal study was compared with that of random samples of school children, who received this task for the first time, substantial practice effects were found. Children of the longitudinal sample not only outperformed the control children regarding strategy use and memory performance, but also showed better task-specific metamemory when asked to explain their behavior. Obviously, it does not require much effort to improve children's strategy knowledge.

Although this finding indicates that children's experiences with memory tasks must play an important role in shaping their use of and knowledge about how to learn and remember, earlier field-studies on the issue were rather disappointing. Teachers did not seem to be able to foster children's metacognitive development in studies that focused on observations of normal classroom situations. For instance, Moely and colleagues observed in classrooms to find out how elementary teachers instructed strategy use and memory knowledge as they presented lessons to children in Grades K to 6 (see the overview by Moely, Santulli, & Obach, 1995). Teachers varied widely in the extent to which they focused on how children might adjust or regulate their cognitive activities in order to master a task. Moely and colleagues concluded from their findings that although explicit instruction to use strategies was rare, a few teachers were quite successful in inducing metacognitive knowledge on memory while working with their elementary school children. Students in classes where teachers employed more strategy suggestions were more likely to engage in spontaneous strategy use. Taken together, the work by Moely and colleagues illustrated that effective teaching processes do not necessarily constitute the rule, and that effective teachers may represent a minority group in elementary school classrooms. The findings indicate that there is a lot of potential for metacognitively guided instructional processes in children's school learning.

More recent research by Ornstein, Coffman, and their colleagues (e.g., Coffman et al., 2008; Grammer et al., 2013; for a review, see Coffman & Cook, 2021) also focused on aspects of the classroom setting, investigating the relation between teachers' memory-related instructions and children's acquisition of memory strategies. Coffman, Ornstein, and colleagues observed that first-grade teachers tend to engage in "memory talk," including strategy suggestions and metacognitive questioning. First-grade children taught by "high mnemonic" teachers who used more of this type of memory talk benefitted more from a memory-strategy training than those children with "low mnemonic" teachers (Coffman et al., 2008). Interestingly enough, the long-term impact of the first-grade teachers was still observed in second grade when children were taught by other teachers (Ornstein, Grammer, & Coffman, 2010). Also, as noted by Grammer et al. (2013), first- and second-graders exposed to memory-rich teaching exhibited greater levels of strategic knowledge and used strategies more efficiently on a memory task involving instructional content than did students exposed to low memory instructions. The longitudinal findings also indicated that the linkage between teachers' "memory talk" and children's development of more complex memory and study skills could still be found in fourth graders (Coffman & Cook, 2021). Thus, this research suggests that "teacher talk" may be relevant for the emergence of mnemonic skills, and that there is a causal link between teachers' language and children's strategy use.

5.2 Metacognition and instruction programs

Strategy training in the classroom. Although rich strategy instruction is not common in schools, it can be successfully implemented. Several comprehensive research projects focused on reading instruction and comprehension monitoring (cf. Borkowski & Mutukrishna, 1995; Paris & Oka, 1986). A particularly important instructional procedure in this context is *reciprocal teaching* (see Palincsar, 1986; Palincsar & Brown, 1984). Reciprocal teaching takes place in a collaborative learning context and involves guided practice in the flexible use of the following four comprehension monitoring strategies: questioning, summarizing, clarifying, and predicting. The novice's role is facilitated by the provision of scaffolding by the expert (teacher). Skills and strategies are practiced in the context of reciprocal teaching dialogs. The teacher and the students take turns leading discussions regarding the contents of a text they are jointly attempting to understand. Overall, this instructional approach has proven to be extraordinarily successful both with normal and learning-disabled students.

Very ambitious studies were undertaken by Pressley and colleagues in order to evaluate effective instructional programs in US public school systems (see Pressley & Hilden, 2006; Schneider & Pressley, 1997). Strategy instruction was not conducted in isolation but was viewed as an integral part of the curriculum, and thus was taught as part of language arts, mathematics, science, and social studies. The goal was to simultaneously enhance children's repertoires of strategies, knowledge, metacognition, and motivation. In accord with the assumption of the Good-Information-Processing-Model outlined above (cf. Pressley et al., 1989), effective teachers did not emphasize the use of single strategies but taught the flexible use of a range of procedures that corresponded to subject matter, time constraints, and other task demands. On most occasions, strategy instruction occurred in groups, with the teachers modeling appropriate strategy use. Pressley and colleagues found that effective teachers regularly incorporated strategy instruction and metacognitive information about flexible strategy use and modification as a part of daily instruction.

Strategy training in addition to classroom lessons. There is evidence that metacognitive competences cannot only be trained by (modeling) teachers' instructional behavior during lessons, but also with short, specialized intervention programs. Wassenburg et al. (2015) conducted an inconsistency-detection training to improve comprehension monitoring in students in Grades 3 and 4. The children participated in the training two times per week over the course of 4 weeks. The training first provided declarative knowledge about comprehension monitoring strategies. The children then trained the correct application of comprehension monitoring strategies. Over the course of the training, the texts, the children worked with became more diverse and challenging. Results on the efficiency of the training showed that comprehension monitoring (measured with an inconsistency task) significantly improved (compared to a control group) in students in Grade 4, who had participated in the program. In students in Grade 3, there was no significant effect.

In another study (Sontag & Stoeger, 2015), an intervention program to train cognitive and metacognitive strategies in the domain of reading comprehension was implemented for students in Grade 4. Within 7 weeks, the students learned cognitive reading strategies, such as underlining main ideas in a text, and metacognitive strategies, such as comparing their own assessment of text comprehension with the number of correctly identified main ideas of the text. The training positively affected the identification of main ideas in expository texts for students at different ability and performance levels.

Domain-specificity of metacognitive competences. Metacognitive knowledge has often been claimed to be context-dependent and domain-specific during an early stage of development, whereas it is supposed to generalize throughout primary school and beyond (Pressley et al., 1989). However, results from studies testing this assumption are inconclusive. In the study by Schneider et al. (2017), there was little evidence for the assumption of an increasingly general character of metacognitive knowledge in the domains of reading comprehension and mathematical competences from Grade 5 to Grade 9: Metacognitive knowledge continued to show a strong domain-specific structure until the end of ninth grade. In contrast, the findings of Veenman and Spaans (2005) indicate that at least in related subjects such as mathematics and biology, metacognitive competences become less domain-specific and more general during a comparable age span.

Overall, research confirms the view that metacognitive knowledge and self-regulated, insightful use of learning strategies not only are influential in elementary school children but are also closely related to reading comprehension in secondary school students. They also give evidence that metacognitive knowledge relevant for school-related domains normally develops during the course of primary school but is not at peak in adolescence. Findings from various intervention approaches show that instructional settings including the training of metacognitive skills can be successful from early school age on, and can still be effective in late childhood and early adolescence (Pressley & Hilden, 2006).

6. Conclusions and implications for future research

The reported theories and empirical findings emphasize that metacognitive competences are an important prerequisite for cognitive performance in a variety of academic domains. These competences develop during childhood and adolescence and can be improved through systematic training. In the domain of memory research, the developmental trajectories of metacognitive competences have been thoroughly explored, while in other domains such as reading comprehension or mathematical competences research has gained momentum in recent years. Open questions remain concerning underlying cognitive functions, for example, executive functioning and working memory that might drive the development of and explain individual differences in both declarative and procedural metacognitive competences. Additionally, promising directions for future research might be comparisons between different indicators of procedural metacognitive competences and the further improvement of online-indicators of such competences (e.g., calibration, detection of inconsistencies, eye-movements). This is particularly relevant when investigating metacognitive competences in different domains and educational contexts.

The origins of metacognition research lie in developmental memory research and thus in the study of metamemory. The impetus for research in the 1970s and 1980s to address metacognitive processes was to understand memory development in children. At the same time, partly inspired by this early work, partly independent of it, the concept of metacognition found its way into other domains. Baker, for example, examined individual differences in metacomprehension as early as the 1980s. In this chapter, based on initial conceptualizations, current studies and research trends were presented exemplarily for the domains of metamemory and metacomprehension, which are particularly relevant for learning in schools. However, a transfer of the concept of metacognition to a variety of other (learning) domains (e.g., mathematics and science, media usage behavior) is also possible and has already been implemented in the past. The general ideas and models of declarative knowledge, monitoring, and regulation can be transferred from metamemory to other domains of metacognitive competences, while content-specific measures and interventions are necessary for each domain. The question of domain specificity of metacognitive competences, also with respect to developmental speed and differentiation, remains to be clarified by future research.

References

Artelt, C., Schiefele, U., & Schneider, W. (2001). Predictors of reading literacy. *European Journal of Psychology of Education*, *16*, 363–383. https://doi.org/10.1007/BF03173188.

Artelt, C., & Schneider, W. (2015). Cross-country generalizability of the role of metacognitive knowledge in students 'strategy use and reading competence'. *Teachers College Record*, *117*(1), 1–32. https://doi.org/10.1177/016146811511700104.

Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory of mind development. *Developmental Psychology*, *35*(5), 1311–1320. https://doi.org/10.1037/0012-1649.35.5.1311.

Baker, L. (1989). Metacognition, comprehension monitoring, and the adult reader. *Educational Psychology Review*, *1*(1), 3–38. https://doi.org/10.1007/BF01326548.

Baker, L., & Anderson, R. I. (1982). Effects of inconsistent information on text processing: Evidence for comprehension monitoring. *Reading Research Quarterly*, *17*(2), 281–294. https://doi.org/10.2307/747487.

Bauer, P. J., & Fivush, R. (Eds.). (2014). *The Wiley handbook on the development of children's memory* Wiley.

Best, D. L., & Folds-Bennett, T. H. (2021). Overview: Deliberate remembering in children. In L. E. Baker-Ward, D. F. Bjorklund, & J. L. Coffman (Eds.), *The development of children's memory—The scientific contributions of Peter A. Ornstein* (pp. 45–61). Cambridge University Press.

Boekaerts, M., & Rozendaal, J. S. (2010). Using multiple calibration indices in order to capture the complex picture of what affects students' accuracy of feeling of confidence. *Learning and Instruction*, *20*(5), 372–382. https://doi.org/10.1016/j.learninstruc.2009.03.002.

Borkowski, J. G., & Mutukrishna, N. (1995). Learning environments and skill generalization: How contexts facilitate regulatory processes and efficacy beliefs. In F. E. Weinert, & W. Schneider (Eds.), *Memory performance and competencies: Issues of growth and development* (pp. 283–300). Lawrence Erlbaum Associates.

Brown, A. L., Bransford, J. D., Ferrara, R. A., & Campione, J. C. (1983). Learning, remembering, and understanding. In J. H. Flavell, & E. M. Markham (Eds.), *Handbook of child psychology: Vol III. Cognitive development* (pp. 77–166). Wiley.

Cavanaugh, J. C., & Perlmutter, M. (1982). Metamemory: A critical examination. *Child Development*, *53*(1), 11–28. https://doi.org/10.2307/1129635.

Coffman, J. L., & Cook, O. K. (2021). Overview: The development of memory in in the elementary classroom context. In L. E. Baker-Ward, D. F. Bjorklund, & J. L. Coffman (Eds.), *The development of children's memory—The scientific contributions of Peter A. Ornstein* (pp. 237–257). Cambridge University Press.

Coffman, J. L., Ornstein, P., McCall, L., & Curran, P. (2008). Linking teachers' memory-relevant language and the development of children's memory skills. *Developmental Psychology*, *44*(6), 1640–1654. https://doi.org/10.1037/a0013859.

Connor, C. M., Radach, R., Vorstius, C., Day, S. L., McLean, L., & Morrison, F. J. (2015). Individual differences in fifth graders' literacy and academic language predict comprehension monitoring development: An eye-movement study. *Scientific Studies of Reading*, *19*(2), 114–134. https://doi.org/10.1080/10888438.2014.943905.

Currie, N. K., Francey, G., Davies, R., Gray, S., Bridges, M. S., Restrepo, M. A., et al. (2020). The process and product of coherence monitoring in young readers: Effects of reader and text characteristics. *Scientific Studies of Reading*, *25*(2), 141–158. https://doi.org/10.1080/10888438.2020.1831503.

de Bruin, A. B. H., Thiede, K. W., Camp, G., & Redford, J. (2011). Generating keywords improves metacomprehension and self-regulation in elementary and middle school children. *Journal of Experimental Child Psychology*, *109*(3), 294–310. https://doi.org/10.1016/j.jecp.2011.02.005.

DeMarie, D., Miller, P. H., Ferron, J., & Cunningham, W. R. (2004). Path analysis tests of theoretical models of children's memory performance. *Journal of Cognition and Development*, *5*(4), 461–492. https://doi.org/10.1207/s15327647jcd0504_4.

Edossa, A. K., Neuenhaus, N., Artelt, C., Lingel, K., & Schneider, W. (2018). Developmental relationship between declarative metacognitive knowledge and reading comprehension during secondary school. *European Journal of Psychology of Education*, *34*, 397–416. https://doi.org/10.1007/s10212-018-0393-x.

Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist*, *13*(4), 277–287. https://doi.org/10.1027/1016-9040.13.4.277.

Efklides, A. (2019). Gifted students and self-regulated learning: The MASRL model and its implications for SRL. *High Ability Studies*, *30*(1–2), 79–102. https://doi.org/10.1080/13598139.2018.1556069.

Ehrlich, M. F., Remond, M., & Tardieu, H. (1999). Processing of anaphoric devices in young skilled and less skilled comprehenders: Differences in metacognitive monitoring. *Reading and Writing*, *11*(1), 29–63. https://doi.org/10.1023/A:1007996502372.

Flavell, J. H. (1971). First discussant's comments: What is memory development the development of? *Human Development*, *14*(4), 272–278. https://doi.org/10.1159/000271221.

Flavell, J. H. (1979). Metacognition and cognitive monitoring. A new area of cognitive-developmental inquiry. *American Psychologist*, *34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906.

Flavell, J. H., Miller, P. H., & Miller, S. A. (2002). *Cognitive development* (4th ed.). Pearson.

Flavell, J. H., & Wellman, H. M. (1977). Metamemory. In R. V. Kail, & J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition* (pp. 3–33). Lawrence Erlbaum Associates.

Fleming, S. M., & Frith, C. (Eds.). (2014). *The cognitive neuroscience of metacognition*. Springer.

Fleur, D. S., Bredeweg, B., & van den Bos, W. (2021). Metacognition: Ideas and insights from neuro- and educational sciences. *Science of Learning*, *13*(6), 1–10. https://doi.org/10.1038/s41539-021-00089-5.

Garner, R. (1987). *Metacognition and reading comprehension*. Ablex Publishing Corporation.

Ghatala, E. S., Levin, J. R., Pressley, M., & Goodwin, D. (1986). A componential analysis of effects of derived and supplied strategy-utility information on children's strategy selections. *Journal of Experimental Child Psychology*, *41*(1), 76–92. https://doi.org/10.1016/0022-0965(86)90052-4.

Ghetti, S., Lyons, K. E., Lazzarin, F., & Cornoldi, C. (2008). The development of metamemory monitoring during retrieval: The case of memory strength and memory absence. *Journal of Experimental Child Psychology*, *99*(3), 157–181. https://doi.org/10.1016/j.jecp.2007.11.001.

Grabe, M., Antes, J., Kahn, H., & Kristjanson, A. (1991). Adult and adolescent readers' comprehension monitoring performance: An investigation of monitoring accuracy and related eye movements. *Contemporary Educational Psychology*, *16*(1), 45–60. https://doi.org/10.1016/0361-476X(91)90005-6.

Grammer, J. K., Coffman, J. L., & Ornstein, P. A. (2013). The effect of teachers' memory-relevant language on children's strategy use and knowledge. *Child Development*, *84*(6), 1989–2002. https://doi.org/10.1111/cdev.12100.

Grammer, J. K., Purtell, K. M., Coffman, J. L., & Ornstein, P. A. (2011). Relations between children's metamemory and strategic performance: Time-varying covariates in early elementary school. *Journal of Experimental Child Psychology*, *108*(1), 139–155. https://doi.org/10.1016/j.jecp.2010.08.001.

Hacker, D. J. (1997). Comprehension monitoring of written discourse across early-to-middle adolescence. *Reading and Writing: An Interdisciplinary Journal*, *9*(3), 207–240. https://doi.org/10.1023/A:1007989901667.

Hasselhorn, M. (1990). The emergence of strategic knowledge activation in categorical clustering during retrieval. *Journal of Experimental Child Psychology*, *50*(1), 59–80. https://doi.org/10.1016/0022-0965(90)90032-4.

Hasselhorn, M. (1995). Beyond production deficiency and utilization inefficiency: Mechanisms of the emergence of strategic categorization in episodic memory tasks. In F. E. Weinert, & W. Schneider (Eds.), *Memory performance and competencies: Issues in growth and development* (pp. 141–159). Lawrence Erlbaum Associates.

Helder, A., Van Leijenhorst, L., & van den Broek, P. (2016). Coherence monitoring by good and poor comprehenders in elementary school: Comparing offline and online measures. *Learning and Individual Differences*, *48*, 17–23. https://doi.org/10.1016/j.lindif.2016.02.008.

Joyner, M. H., & Kurtz-Costes, B. (1997). Metamemory development. In N. Cowan (Ed.), *The development of memory in childhood* (pp. 275–300). Psychology Press.

Koriat, A. (2008). Easy comes, easy goes? The link between learning and remembering and its exploitation in metacognition. *Memory & Cognition*, *36*, 416–428. https://doi.org/10.3758/MC.36.2.416.

Koriat, A., Ackerman, R., Adiv, S., Lockl, K., & Schneider, W. (2014). The effects of goal-driven and data-driven regulation on metacognitive monitoring during learning: A developmental perspective. *Journal of Experimental Psychology: General*, *143*(1), 386–403. https://doi.org/10.1037/a0031768.

Koriat, A., Ackerman, R., Lockl, K., & Schneider, W. (2009). The memorizing effort heuristic in judgments of learning: A developmental perspective. *Journal of Experimental Child Psychology*, *102*(3), 265–279. https://doi.org/10.1016/j.jecp.2008.10.005.

Kreutzer, M. A., Leonard, C., & Flavell, J. H. (1975). An interview study of children's knowledge about memory. *Monographs of the Society for Research in Child Development*, *40*(1), 1–60. https://doi.org/10.2307/1165955.

Kron-Sperl, V., Schneider, W., & Hasselhorn, M. (2008). The development and effectiveness of memory strategies in kindergarten and elementary school: Findings from the Würzburg and Göttingen longitudinal memory studies. *Cognitive Development*, *23*(1), 79–104. https://doi.org/10.1016/j.cogdev.2007.08.011.

Kuhn, D. (2000). Theory of mind, metacognition, and reasoning: A life-span perspective. In P. Mitchell, & K. J. Riggs (Eds.), *Children's reasoning and the mind* (pp. 301–326). Psychology Press.

Lingel, K., Lenhart, J., & Schneider, W. (2019). Metacognition in mathematics: Do different metacognitive monitoring measures make a differerence? *ZDM Mathematics Education*, *51*(4), 587–600. https://doi.org/10.1007/s11858-019-01062-8.

Lipowski, S. L., Merriman, W. E., & Dunlosky, J. (2013). Preschoolers can make highly accurate judgments of learning. *Developmental Psychology*, *49*(8), 1505–1516. https://doi.org/10.1037/a0030614.

Lockl, K., & Schneider, W. (2002). Developmental trends in children's feeling-of-knowing judgements. *International Journal of Behavioral Development*, *26*(4), 327–333. https://doi.org/10.1080/01650250143000210.

Lockl, K., & Schneider, W. (2007). Knowledge about the mind: Links between theory of mind and later metamemory. *Child Development*, *78*(1), 148–167. https://doi.org/10.1111/j.1467-8624.2007.00990.x.

Long, D. L., & Chong, J. L. (2001). Comprehension skill and global coherence: A paradoxical picture of poor comprehenders' abilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(6), 1424. https://doi.org/10.1037/0278-7393.27.6.1424.

Moely, B. E., Santulli, K. A., & Obach, M. S. (1995). Strategy instruction, metacognition, and motivation in the elementary school classroom. In F. E. Weinert, & W. Schneider (Eds.), *Memory performance and competencies: Issues in growth and development* (pp. 301–321). Lawrence Erlbaum Associates.

Münchow, H., Richter, T., von der Mühlen, S., & Schmid, S. (2019). The ability to evaluate arguments in scientific texts: Measurement, cognitive processes, nomological network and relevance for academic success at the university. *British Journal of Educational Psychology*, *89*(3), 502–523. https://doi.org/10.1111/bjep.12298.

Myers, J. L., & O'Brien, E. J. (1998). Accessing the discourse representation during reading. *Discourse Processes*, *26*(2–3), 131–157. https://doi.org/10.1080/01638539809545042.

Myers, M., & Paris, S. G. (1978). Children's metacognitive knowledge about reading. *Journal of Educational Psychology*, *70*(5), 680–690. https://doi.org/10.1037/0022-0663.70.5.680.

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe, & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–25). MIT Press.

Oakhill, J. V., & Cain, K. (2012). The precursors of reading comprehension and word reading in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading*, *16*(2), 91–121. https://doi.org/10.1080/10888438.2010.529219.

Oakhill, J. V., Hartt, J., & Samols, D. (2005). Levels of comprehension monitoring and working memory in good and poor comprehenders. *Reading and Writing*, *18*(7–9), 657–713. https://doi.org/10.1007/s11145-005-3355-z.

O'Brien, E. J., Rizzella, M. L., Albrecht, J. E., & Halleran, J. G. (1998). Updating a situation model: A memory-based text processing view. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(5), 1200–1210. https://doi.org/10.1037/0278-7393.24.5.1200.

Ornstein, P. A., Grammer, J. K., & Coffman, J. L. (2010). Teachers' "mnemonic style" and the development of skilled memory. In H. S. Waters, & W. Schneider (Eds.), *Metacognition, strategy use, and instruction; metacognition, strategy use, and instruction* (pp. 23–53). Guilford Press.

Palincsar, A. S. (1986). The role of dialogue in providing scaffolded instruction. *Educational Psychologist*, *21*(1–2), 73–98. https://doi.org/10.1080/00461520.1986.9653025.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-monitoring activities. *Cognition and Instruction*, *1*(2), 117–175. https://doi.org/10.1207/s1532690xci0102_1.

Paris, S. G., & Oka, E. R. (1986). Children's reading strategies, metacognition, and motivation. *Developmental Review*, *6*(1), 25–56. https://doi.org/10.1016/0273-2297(86)90002-X.

Perner, J. (1991). *Understanding the representational mind*. MIT Press.

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Lawrence Erlbaum Associates.

Pressley, M., Borkowski, J. G., & Schneider, W. (1989). Good information processing: What it is and what education can do to promote it. *International Journal of Educational Research*, *13*(8), 857–867. https://doi.org/10.1016/0883-0355(89)90069-4.

Pressley, M., & Hilden, K. (2006). Cognitive strategies. In D. Kuhn, R. S. Siegler, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Cognition, perception, and language* (pp. 511–556). Wiley.

Richter, T. (2015). Validation and comprehension of text information: Two sides of the same coin. *Discourse Processes*, *52*(5–6), 337–355. https://doi.org/10.1080/0163853X.2015.1025665.

Roebers, C. M. (2002). Confidence judgments in children's and adult's event recall and suggestibility. *Developmental Psychology*, *38*(6), 1052–1067. https://doi.org/10.1037/0012-1649.38.6.1052.

Roebers, C. M. (2014). Children's deliberate memory development: The contribution of strategies and metacognitive processes. In P. J. Bauer, & R. Fivush (Eds.), *Vol. II. The Wiley handbook on the development of children's memory* (pp. 865–894). Wiley.

Roebers, C., von der Linden, N., Howie, P., & Schneider, W. (2007). Children's metamemorial judgments in an event recall task. *Journal of Experimental Child Psychology*, *97*(2), 117–137. https://doi.org/10.1016/j.jecp.2006.12.006.

Savolainen, H., Ahonen, T., Aro, M., Tolvanen, A., & Holopainen, L. (2008). Reading comprehension, word reading and spelling as predictors of school achievement and choice of secondary education. *Learning and Instruction*, *18*(2), 201–210. https://doi.org/10.1016/j.learninstruc.2007.09.017.

Schlagmüller, M., & Schneider, W. (2002). The development of organizational strategies in children: Evidence from a microgenetic longitudinal study. *Journal of Experimental Child Psychology*, *81*(3), 298–319. https://doi.org/10.1006/jecp.2002.2655.

Schlagmüller, M., & Schneider, W. (2007). *WLST 7-12. Würzburger Lesestrategie-Wissenstest für die Klassen 7–12 [Würzburg reading strategy knowledge test for grades 7–12]*. Hogrefe.

Schneider, W. (2010). Metacognition and memory development in childhood and adolescence. In H. S. Waters, & W. Schneider (Eds.), *Metacognition, strategy use, and instruction* (pp. 54–81). Guilford Press.

Schneider, W. (2015). *Memory development from early childhood through emerging adulthood*. Springer.

Schneider, W., & Bjorklund, D. F. (1998). Memory. In W. Damon, D. Kuhn, & R. S. Siegler (Eds.), *Handbook of child psychology, Vol. 2. Cognition, perception, and language* (5th ed., pp. 467–521). Wiley.

Schneider, W., Kron, V., Hünnerkopf, M., & Krajewski, K. (2004). The development of young children's memory strategies: Findings from the Würzburg longitudinal memory study. *Journal of Experimental Child Psychology*, *88*(2), 193–209. https://doi.org/10.1016/j.jecp.2004.02.004.

Schneider, W., Lingel, K., Artelt, C., & Neuenhaus, N. (2017). Metacognitive knowledge in secondary school students: Assessment, structure, and developmental change. In D. Leutner, J. Fleischer, J. Grünkorn, & E. Klieme (Eds.), *Competence assessment in education: Research, models, and instruments* (pp. 285–302). Springer.

Schneider, W., & Lockl, K. (2002). The development of metacognitive knowledge in children and adolescents. In T. J. Perfect, & B. L. Schwartz (Eds.), *Applied metacognition* (pp. 224–257). Cambridge University Press.

Schneider, W., & Lockl, K. (2008). Procedural metacognition in children: Evidence for developmental trends. In J. Dunlosky, & R. A. Bjork (Eds.), *Handbook of metamemory and memory* (pp. 391–409). Routledge.

Schneider, W., & Löffler, E. (2016). The development of metacognitive knowledge in children and adolescents. In J. Dunlosky, & S. K. Tauber (Eds.), *The Oxford handbook of metamemory* (pp. 491–518). Oxford University Press.

Schneider, W., & Ornstein, P. A. (2019). Determinants of memory development in childhood and adolescence. *International Journal of Psychology*, *54*(3), 307–315. https://doi.org/10.1002/ijop.12503.

Schneider, W., & Pressley, M. (1997). *Memory development between 2 and 20*. Lawrence Erlbaum Associates.

Schneider, W., Schlagmüller, M., & Visé, M. (1998). The impact of metamemory and domain-specific knowledge on memory performance. *European Journal of Psychology of Education*, *13*(1), 91–103. https://doi.org/10.1007/BF03172815.

Schneider, W., Visé, M., Lockl, K., & Nelson, T. O. (2000). Developmental trends in children's memory monitoring: Evidence from a judgment-of-learning task. *Cognitive Development*, *15*(2), 115–134. https://doi.org/10.1016/S0885-2014(00)00024-1.

Schraw, G., Kuch, F., Gutierrez, A. P., & Richmond, A. S. (2014). Exploring a three-level model of calibration accuracy. *Journal of Educational Psychology*, *106*(4), 1192–1202. https://doi.org/10.1037/a0036653.

Schraw, G., & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review*, *7*(4), 351–371. https://doi.org/10.1007/BF02212307.

Sodian, B., & Schneider, W. (1999). Memory strategy development—Gradual increase, sudden insight or roller coaster? In F. E. Weinert, & W. Schneider (Eds.), *Individual* *development from 3 to 12: Findings from the Munich longitudinal study* (pp. 61–77). Cambridge University Press.

Son, L. K., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(1), 204–221. https://doi.org/10.1037/0278-7393.26.1.204.

Sontag, C., & Stoeger, H. (2015). Can highly intelligent and high-achieving students benefit from training in self-regulated learning in a regular classroom context? *Learning and Individual Differences*, *41*, 43–53. https://doi.org/10.1016/j.lindif.2015.07.008.

Tibken, C., Richter, T., von der Linden, N., Schmiedeler, S., & Schneider, W. (2022). The role of metacognitive competences in the development of school achievement among gifted adolescents. *Child Development*, *93*(1), 117–133. https://doi.org/10.1111/cdev.13640.

Tibken, C., Richter, T., Wannagat, W., Schmiedeler, S., von der Linden, N., & Schneider, W. (2022). Measuring comprehension monitoring with the inconsistency task in adolescents: Stability, associations with reading comprehension skills, and differences between grade levels. *Discourse Processes*. Advance online publication. https://doi.org/10.1080/0163853X.2022.2073736 (open access).

van den Broek, P., & Helder, A. (2017). Cognitive processes in discourse comprehension: Passive processes, reader-initiated processes, and evolving mental representations. *Discourse Processes*, *54*(5–6), 360–372. https://doi.org/10.1080/0163853X.2017.1306677.

van den Broek, P., Risden, K., & Husebye-Hartman, E. (1995). The role of readers' standards for coherence in the generation of inferences during reading. In R. F. Lorch, & E. J. O'Brien (Eds.), *Sources of coherence in reading* (pp. 353–373). Lawrence Erlbaum Associates.

van der Schoot, M., Reijntjes, A., & van Lieshout, E. C. D. M. (2012). How do children deal with inconsistencies in text? An eye fixation and self-paced reading study in good and poor reading comprehenders. *Reading and Writing: An Interdisciplinary Journal*, *25*(7), 1665–1690. https://doi.org/10.1007/s11145-011-9337-4.

van Moort, M. L., Koornneef, A., & van den Broek, P. W. (2021). Differentiating text-based and knowledge-based validation processes during reading: Evidence from eye movements. *Discourse Processes*, *58*(1), 22–41. https://doi.org/10.1080/0163853X.2020.1727683.

Veenman, M. V. J., & Spaans, M. A. (2005). Relation between intellectual and metacognitive skills: Age and task differences. *Learning and Individual Differences*, *15*(2), 159–176. https://doi.org/10.1016/j.lindif.2004.12.001.

von der Linden, N., Löffler, E., & Schneider, W. (2017). Relation of metacognitive monitoring and control processes across the life span. *Journal of Educational and Developmental Psychology*, 7(1), 86–101. https://doi.org/10.5539/jedp.v7n1p86.

von der Linden, N., & Roebers, C. M. (2006). Developmental changes in uncertainty monitoring during an event recall task. *Metacognition and Learning*, *1*, 213–228. https://doi.org/10.1007/s11409-006-9001-6.

Wassenburg, S. I., Bos, L. T., de Koning, B. B., & van der Schoot, M. (2015). Effects of an inconsistency-detection training aimed at improving comprehension monitoring in primary school children. *Discourse Processes*, *52*(5–6), 463–488. https://doi.org/10.1080/0163853X.2015.1025203.

Waters, H. S., & Kunnmann, T. (2010). Metacognition and strategy discovery in early childhood. In H. S. Waters, & W. Schneider (Eds.), *Metacognition, strategy use, and instruction* (pp. 3–22). Guilford Press.

Wellman, H. M. (1978). Knowledge of the interaction of memory variables: A developmental study of metamemory. *Developmental Psychology*, *14*, 24–29.

Wellman, H. M. (1990). *The child's theory of mind*. MIT Press.

Zabrucky, K., & Ratner, H. H. (1986). Children's comprehension monitoring and recall of inconsistent stories. *Child Development*, *57*(6), 1401–1418. https://doi.org/10.2307/1130419.

Zabrucky, K., & Ratner, H. H. (1989). Effects of reading ability on children's comprehension evaluation and regulation. *Journal of Reading Behavior*, *21*(1), 69–83. https://doi.org/10.1080/10862968909547659.

Zabrucky, K., & Ratner, H. H. (1992). Effects of passage type on comprehension monitoring and recall in good and poor readers. *Journal of Reading Behavior*, *24*(3), 373–391.

Zargar, E., Adams, A. M., & Connor, C. M. (2019). The relations between children's comprehension monitoring and their reading comprehension and vocabulary knowledge: An eye-movement study. *Reading and Writing: An Interdisciplinary Journal*, *33*(4), 511–545. https://doi.org/10.1007/s11145-019-09966-3.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*(2), 162–185. https://doi.org/10.1037/0033-2909.123.2.162.

Further reading

Schunk, D. H., & Zimmerman, B. J. (Eds.). (1998). *Self-regulated learning: From teaching to self-reflective practice* Guilford Press.

Veenman, M., Kok, R., & Blöte, A. (2005). The relation between intellectual and metacognitive skills in early adolescence. *Instructional Science*, *33*, 193–211. https://doi.org/10.1007/s11251-004-2274-8.

This page intentionally left blank

CHAPTER ELEVEN

Learning about others and learning from others: Bayesian probabilistic models of intuitive psychology and social learning

Rongzhi Liu* **and Fei Xu**
Department of Psychology, University of California, Berkeley, CA, United States
*Corresponding author: e-mail address: rongzhi_liu@berkeley.edu

Contents

Abstract

How do infants and young children reason about other people? What inferences do they make when they learn from teachers and whom do they choose to learn from? Past research in developmental psychology has demonstrated infants' and young children's competence in making these inferences. However, the mechanisms underlying these inferences and how these mechanisms change across development are less clear. In this chapter, we review a growing body of Bayesian probabilistic models on intuitive psychology and social learning. We integrate these models with past and new empirical studies within the framework of rational constructivism. These models showed that infants and children have intuitive theories about others (agents, teachers, and informants). When given new evidence, they rationally update their beliefs about others and their beliefs about the world based on these intuitive theories. Developmental changes can be explained by advances in children's intuitive theories. Finally, we propose future directions for both empirical and modeling work in these domains.

Advances in Child Development and Behavior, Volume 63
ISSN 0065-2407
https://doi.org/10.1016/bs.acdb.2022.04.007

1. Introduction

Infants are born into a world full of other individuals. From the beginning of life, infants are interested in other people (e.g., newborns prefer looking at human faces; Johnson, Dziurawiec, Ellis, & Morton, 1991). As they observe and interact with them, infants gradually come to understand other individuals. They learn that individuals' actions are driven by unobservable and distinct mental states. They make inferences about others' mental states from their actions. To what extent are these reasoning and inferences rational?

Other people are also an important source to learn from. Humans possess powerful social learning abilities that allow the accumulation of rich knowledge and skills over generations (Boyd, Richerson, & Henrich, 2011; Tomasello, 2016). What kinds of mechanisms underlie our abilities to learn from other people and teach other people? How do we decide whom to learn from? To what extent are these inferences rational?

The purpose of this chapter is to review a growing body of Bayesian probabilistic models on intuitive psychology and social learning that can shed light on the development of these abilities. We will situate our discussion of these models in the *rational constructivism* theory of cognitive development (Gopnik & Wellman, 2012; Xu, 2019; Xu & Griffiths, 2011; see Ullman & Tenenbaum, 2020, for a related perspective on learning as hierarchical Bayesian program induction). Rational constructivism characterizes infants' initial representations as a set of proto-conceptual primitives and the mature state of human conceptual system as a set of domain-specific intuitive theories. Most importantly, a key mechanism that drives learning from the initial state to the mature state is the rational statistical inference that underlies Bayesian probabilistic models. This theoretical framework has been successful in bringing together computational modeling and empirical work to shed light on a variety of topics in cognitive development (Xu & Griffiths, 2011; Xu & Kushnir, 2012). In this chapter, we will integrate recent models on intuitive psychology and social learning with past and new empirical studies on these topics within the framework of rational constructivism. We will highlight new insights provided by the models, and propose future directions that would advance our understanding of intuitive psychology and social learning.

2. Reasoning about others' mental states and actions

As children interact with and learn about the social world, one big challenge is to learn about and understand other individuals. People differ from objects in that each individual has unobservable and distinct mental states (e.g., goals, preferences, beliefs, motivations, emotions) that could only be inferred through behaviors. In developmental psychology, there is a long tradition of studying infants' and children's abilities to reason about others' mental states and actions. Past research has identified several key phenomena in the domain of intuitive psychology.

Infants interpret human's reaching and grasping motions as goal-directed actions (Woodward, 1998). In these studies, 6-month-old infants observed an agent repeatedly reach for and grasp one of two toys. Then, the two toys switched locations. Infants expected the agent to reach for the object she previously selected, which is now at a new location, instead of reaching for the new object at the original location. Thus, infants encoded the selected object as the agent's goal, and expected agents to perform actions consistent with the goal.

Infants expect agents to use the most efficient means to achieve their goals (i.e., the rationality principle), and use this principle to guide their inferences about agents' actions, goals, and situational constraints (Gergely & Csibra, 2003). Nine- and 12-month-olds saw simple animations where a small circle (the actor) reached a large circle (the goal) by jumping over a wall separating them. When the wall was removed, infants expected the agent to take a novel but efficient, straight-line path to reach the goal, instead of the same, jumping path (Csibra, Gergely, Bíró, Koós, & Brockbank, 1999; Gergely, Nádasdy, Csibra, & Bíró, 1995). Twelve-month-old infants can infer agents' goals even when they have not seen the goals being achieved (Csibra, Bíró, Koós, & Gergely, 2003). They observed a large circle approaching a moving small circle as if it was chasing the small circle. Then, the small circle went through a small gap between two walls. The large circle could not go through the gap, so it went around the walls. When the small circle stopped, 12-month-olds expected the large circle to approach and make contact with the small circle, rather than pass by the small circle and continue moving. In other words, they attributed the goal of catching the small circle to the large circle even though they had not seen the large circle achieve this goal. Lastly, 12-month-olds can infer unobserved constraints from agents' actions and goals (Csibra et al., 2003). A screen occluded the view between a small circle (the actor) and a large circle (the goal). Infants observed the small circle taking a jumping path to reach the large circle. When the screen was removed, infants expected that there would be a wall (a constraint) between the small circle and the large circle.

Infants and young children can also infer others' preferences from sampling behaviors (Kushnir, Xu, & Wellman, 2010; Wellman, Kushnir, Xu, & Brink, 2016). For instance, 10-month-olds observed an agent sampled 5 target objects from a jar that consisted of either 20% of target objects or 80% of target objects. Thus, when the jar consisted of 20% of target objects, the agent violated random sampling. But when the jar consisted of 80% of target objects, the agent did not violate random sampling. Infants only inferred that the agent preferred the target objects and expected the agent to choose the target object again in the 20% condition (Wellman et al., 2016).

Lastly, infants socially evaluate agents based on their helping or hindering behaviors toward third parties (Hamlin & Wynn, 2011; Hamlin, Wynn, & Bloom, 2007). For instance, 6- and 10-month-olds observed the interactions among three agents (depicted as geometric shapes with eyes). The climber attempted to climb a hill. On alternating trials, infants observed a helper pushing the climber up, and a hinderer pushing the climber down. Infants showed a preference for the helper agent over the hinderer agent—they were more likely to reach for the helper agent than the hinderer agent when they were given a choice (Hamlin et al., 2007).

What is the computational basis of these fundamental inferences about agents' mental states and actions? Can these inferences be considered rational? In recent years, there has been a surge in using Bayesian probabilistic models to capture these inferences. These models use generative models (formalized intuitive theories) to specify how agents plan their actions based on their mental states, and then use Bayesian inference over the generative models to infer unobservable mental states from observable actions. In the following sections, we review three categories of Bayesian models on early intuitive psychology: inverse planning models, inverse decision-making models, and the naïve utility calculus. For each category of models, we will first lay out the technical details of a representative model from that category, and then describe how these models have been applied to capture the key intuitive psychology phenomena observed in infants and children.

2.1 Inverse planning models

Inverse planning models capture the joint inferences of a variety of mental states—goals, beliefs, desires—from agents' actions and the constraints in the environments. This approach uses a generative model to capture agents' action planning process, assuming that observers represent other agents as rational planners. The generative model is then inverted using Bayesian inference to infer the agent's mental states based on observed actions.

A representative inverse planning model is the Bayesian theory of mind (BToM) model developed by Baker, Jara-Ettinger, Saxe, and Tenenbaum (2017). BToM aims to reverse-engineer the elementary form of mental state inferences that emerges in infancy, including goal encoding (Woodward, 1998), teleological representation of actions (Csibra et al., 2003), false belief understanding (Onishi and Baillargeon, 2005), and social evaluation (Hamlin et al., 2007). BToM formalizes these mental state inferences as Bayesian inference over a generative model of a rational agent. The generative model uses partially observable Markov decision processes (POMDPs)—an artificial intelligence approach to rational planning and state estimation.

We will use a simple scenario in Fig. 1 to illustrate the model details. A student wants to buy lunch from a food truck. There are three types of food trucks, Korean (K), Lebanese (L), and Mexican (M), and two parking spaces (the yellow squares). The blue circle represents the student, and the black dots represent the traces of the student's movements. Unshaded area in the scenes is perceptually accessible to the student; shaded area is not perceptually accessible to the student. In this example, the student started at a place where she saw the Korean food truck was parked at the bottom-left parking space, but she could not see the top-right parking space. She moved toward a place where she saw that the Lebanese food truck was parked at the top-right parking space. Then, she returned to the bottom-left parking space and bought Korean food.

Fig. 1 Frames of the food truck scenario in Baker et al. (2017).

In the model, the agent's beliefs are represented as a probability distribution over all possible world states. In the food truck scenario, world states include all the possible combinations of the food trucks that came to campus and the parking locations of each food truck. The agent's beliefs can be updated given the agent's new percepts. For instance, in frame 2 (the top-right frame) of Fig. 1, the agent is able to see the parking space in the top-right corner, and her beliefs are updated. The agent's desires are represented by a utility function over situations, actions, and events. In the food truck scenario, the agent's desires are her utilities for eating at each food truck. BToM starts with prior beliefs about the agent's beliefs and desires, and jointly infers the posterior probability of the agent's beliefs (B), desires (D), percepts (P) and the situations (S) given the agent's actions (A). Formally, the posterior distribution is proportional to the product of the prior distribution and three likelihood terms:

$$\Pr(B, D, P, S \mid A) \propto \Pr(A \mid B_1, D) \times \Pr(B_1 \mid P, B_0) \times \Pr(P \mid S) \times \Pr(B_0, D, S).$$

$\Pr(B_0, D, S)$ corresponds to the prior distribution of the agent's belief, desire, and the situation. $\Pr(P \mid S)$ corresponds to how the agent forms her percepts given the situation; $\Pr(B_1 \mid P, B_0)$ corresponds to how the agent updates her belief given the initial belief and new percepts; $\Pr(A \mid B_1, D)$ corresponds to how the agent plans her actions given her beliefs and desires. For the last likelihood term, the model assumes that the agent will achieve her desires by choosing the most efficient actions (the principle of efficiency). Since an agent's behaviors might deviate from the rational model, the model adopts a graded expectation of utility maximization: the agent is most likely to choose the highest-utility action at each step in the planning process, but she sometimes chooses a non-optimal action.

Baker et al. (2017) also developed a few alternative models. Two of them are lesioned versions of the full BToM model. One lesioned model does not represent uncertain beliefs; it assumes that agents' beliefs are always the same as the true world state. The other lesioned model assumes that agents' actions do not involve costs; therefore, the model does not incorporate the principle of efficiency. A third alternative model is a motion-based heuristic model, which makes predictions based on the learned statistical associations between motion and environmental cues and people's judgments about agents' mental states. The performances of BToM and the alternative models were compared to adults' performances in two types of tasks. The first type of task involves inferring an agent's desires and beliefs from the agent's actions and the environment. After observing scenarios similar to the one in Fig. 1, adults and the models inferred the agent's preferences for the three food trucks and her initial beliefs about which food truck was parked at the unobservable space. In the second type of task, adults and models inferred the states of the world from an agent's actions and desires. They observed scenarios where the agent moved in a complex environment, searching for his favorite food truck (known to the participants). They inferred the locations of the food trucks in the environment based on the agent's actions. In both types of tasks, BToM performed closer to adults, and better than the lesioned BToM models and the motion-based heuristic model. This suggests that adults' abilities to make joint inferences of beliefs, desires, and percepts in these tasks cannot be achieved by merely learning the statistical associations between motion cues and mental states. Instead, a rational planning model that uses the principle of efficiency and represents uncertain beliefs is underlying adults' mental states inferences.

BToM has only been evaluated in tasks that are more complex than the intuitive psychology phenomena observed in infants. In contrast, Shu et al. (2021) evaluated models on tasks similar to the ones that infants have been tested on. In particular, they focused on four key phenomena: inferring goals from agents' actions (Woodward, 1998), predicting agents' actions based on goals and constraints (Gergely et al., 1995), inferring constraints from agents' actions and goals (Csibra et al., 2003), and inferring agents' preferences from the levels of costs they incurred (Liu, Ullman, Tenenbaum, & Spelke, 2017). Shu et al. (2021) tested the performance of two models on these tasks: an inverse planning model, the Bayesian Inverse Planning and Core Knowledge (BIPaCK), and a neural network model, the Theory of Mind Neural Network (ToMnet-G). BIPaCK is based on a generative model that integrates two components—a physics simulation that depends on core knowledge of objects and physics, and an agent planning process that depends on utility computation (maximizing rewards and minimizing costs). The physics simulation extracts different types of entities from the video—the agent, the goal objects, and obstacles—and recreates an approximated physical scene. The agent planning process is simpler than that in BToM. The model only represents one type of mental state—the agent's goals. The model predicts a trajectory that allows the agent to reach the goal by maximizing the agent's reward and minimizing the agent's cost. Training videos were used to calibrate parameters, and test videos were used to evaluate model performance and generalization. For each test video, the model yields a surprise rating that is defined by the expected distance between the predicted agent trajectory and the one observed in the test video. Model performances were compared to adults' surprise ratings of the test videos.

BIPaCK performed closer to adults than ToMnet-G did, and achieved better performance in generalization within and across scenarios. The findings suggest that an intuitive theory that represents agents' goals and uses utility computation to plan actions underlies infants' reasoning in the four key intuitive psychology phenomena. In addition, core knowledge of objects and physics is an important requirement for reasoning about agents' mental states and actions.

Hamlin, Ullman, Tenenbaum, Goodman, and Baker (2013) used an inverse-planning model to capture infants' social evaluation of agents who help or harm third parties. They used a combination of modeling and behavioral experiments to show that infants' social evaluation is based on inferences about agents' mental states. The study aimed to rule out alternative accounts including the low-level cue-based account (e.g., infants prefer those who push things uphill vs. downhill) and mid-level accounts (e.g., infants only represent the protagonist's first-order goals, and positively evaluate any individuals who complete them; infants only represent second-order goals of helping and harming, but do not understand that helping and harming requires having knowledge about the protagonist's goals). In the experiment, infants observed two Lifter puppets lifting one of two doors, allowing the Protagonist puppet to reach the object behind the doors. One Lifter always allowed the Protagonist to reach the object it had repeatedly grasped before, and the other Lifter always allowed the Protagonist to reach the object it had not grasped before. The experiment varied whether the Protagonist showed preference or not (the Protagonist repeatedly chose an object from two options or only one option), whether the Lifters were knowledgeable about the Protagonist's preference (the Lifters were present or absent when the Protagonist showed preference). The Full Mental model infers the Lifter's goals (prosocial, neutral, or antisocial) based on the Lifter's beliefs and actions, as well as the protagonist's goals. Consistent with the Full Mental model, infants only positively evaluated the Lifter who allowed the Protagonist to reach the preferred object, when the protagonist showed a preference, and when the Lifters are knowledgeable about the preference. The Full Mental model provided a better qualitative fit to infants' behaviors compared to a feature-based model, models that evaluate agents based on completion of the protagonist's first-order goal, or a model that evaluates the second-order goals but not second-order beliefs. Thus, 10-month-olds' social evaluation is based on inferences about the mental states of the protagonist as well as the helper and hinderer.

These models provided new insights into our understanding of the development of intuitive psychology. Previous experiments have demonstrated that even infants can make inferences about others' goals, beliefs, and desires. However, the underlying mechanisms of these inferences were less clear. Some researchers have argued that cue-based and heuristic-based learning can account for infants' and adults' reasoning about agents (Blythe, Todd, & Miller, 1999; Gao, Newman, & Scholl, 2009; Perner & Ruffman, 2005; Scholl & Tremoulet, 2000). For instance, Perner and Ruffman (2005) argued that infants' understanding of false belief in Onishi and Baillargeon (2005) might be achieved by learning the association between agents, objects, and location in the scene. In contrast, the inverse planning models assume that infants, children, and adults have intuitive theories of how agents' mental states relate to their actions, and can make inferences about their mental states based on these intuitive theories. By comparing Bayesian inverse planning models with alternative models, Baker et al. (2017), Shu et al. (2021), and Hamlin et al. (2013) showed that adults' and infants' abilities to reason about agents cannot be achieved by merely learning the statistical associations between motion cues and mental states. Instead, it depends on intuitive theories of agent planning. Both adults and infants have complex intuitive theories that allow joint inferences over various types of mental states (Baker et al., 2017; Hamlin et al., 2013). In addition, Shu et al. (2021) revealed that core knowledge of objects and physics is required for reasoning about agents' mental states and actions. This is consistent with the developmental trajectory of infants' abilities to reason about mental states (emerge around 6–12 months of age; Gergely & Csibra, 2003; Wellman et al., 2016; Woodward, 1998) and their abilities to reason about objects (emerge around 2–6 months of age; Aguiar & Baillargeon, 1999; Leslie & Keeble, 1987; Spelke, Breinlinger, Macomber, & Jacobson, 1992). Based on the findings of Shu et al. (2021), it is possible that mental states inferences emerge slightly later in development because these inferences depend on core knowledge of objects.

An important future direction is to use inverse planning models to examine infants' abilities to understand false beliefs. Given the controversy in the interpretation of the findings in Onishi and Baillargeon (2005), it would be helpful to combine modeling and behavioral approaches to reveal whether the ability to represent false belief is indeed necessary for infants' performance in this task. For instance, this could be achieved by comparing infants' performance in this task to the performance of BToM, a lesioned version of BToM that always represents agents' beliefs as the true world states, and a motion-based heuristic model.

Future works can also extend the inverse planning model of social evaluation to capture more nuanced forms of social evaluation. For instance, Hamlin, Wynn, Bloom, and Mahajan (2011) found that 8-month-olds evaluated agents who showed prosocial or antisocial behaviors based on the dispositional status of the recipient. They preferred an agent who acted prosocially toward a recipient who previously engaged in prosocial behaviors, but they preferred an agent who acted antisocially toward a recipient who previously engaged in antisocial behaviors. However, 5-month-olds did not engage in such nuanced social evaluation—they always preferred a prosocial agent, regardless of the dispositional status of the recipient. The inverse planning model can be used to reveal the generative model (the intuitive theory) underlying these nuanced forms of social evaluation, and to capture the developmental changes between 5- and 8-month-olds.

2.2 Inverse decision-making models

The second category of models on intuitive psychology is inverse decision-making models. These models capture the inferences of a particular type of mental state—preferences. An inverse decision-making model incorporates a decision-making model as the generative model, which specifies how people make choices based on their preferences. Then, the model uses Bayesian inference to invert the decision-making model, and infers agents' preferences from their choice patterns.

Lucas et al. (2014) adopted an econometrics model, the Mixed Multinomial Logit model (MML), to capture the developmental data on children's inferences about others' preferences and choices. The MML specifies a choice model that maps people's preferences to choices. The choice model assumes that people combine the subjective utilities of different features of each option, and choose the option that maximizes their utility. The choice rule specifies that the probability that an agent will choose an option increases exponentially with that option's utility. Then, the MML uses Bayesian inference to invert the choice model, and infer others' preferences based on their choices.

The model captured the developmental data on children's preference inferences. In Kushnir et al. (2010), 20-month-olds and preschoolers observed an agent picking out four target objects from a box of different types of objects. Target objects contained 100%, 50%, or 18% of all objects in the box (Fig. 2). Children were asked to infer the agent's preferred object from the target object, the alternative object in the box, or a novel object.

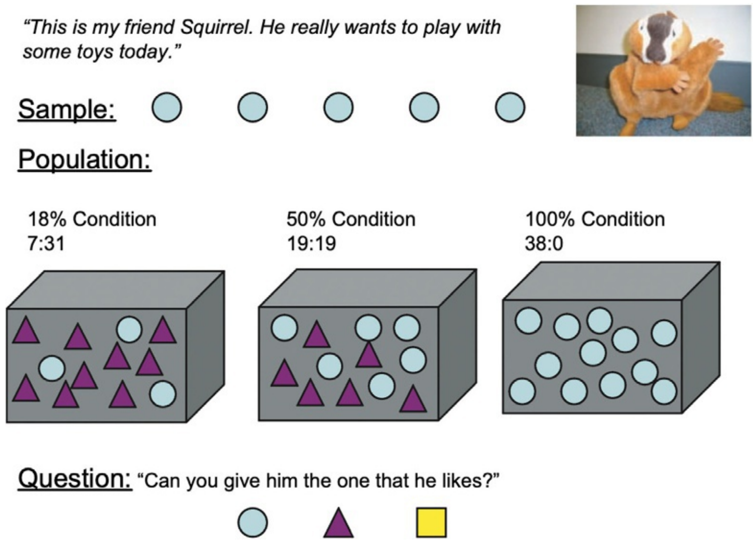


Fig. 2 A schematic depiction of the experimental procedure in Kushnir et al. (2010).

The posterior probability of the agent's preference for each type of object is calculated based on the prior probability of its preference, and the likelihood that the agent will make the observed choices given that preference. In the 100% condition, since the agent could only choose the target objects, his preference would not influence his choices. Thus, the posterior probability of the agent's preference distribution is uninformative. In the 50% condition, the posterior probability would shift toward having a preference for the target object, since the likelihood of observing these choices is low (0.5^5) if the agent had no preference and was sampling randomly. In the 18% condition, the posterior probability will shift toward an even stronger preference for the target object, since the likelihood of observing these choices is even lower (only 0.18^5) under random sampling. Consistent with the model predictions, children chose randomly from the three objects in the 100% condition, and they were increasingly more likely to choose the target object in the 50% and the 18% conditions.

The model also captured the developmental difference in preference understanding observed in Repacholi and Gopnik (1997). In the study, 14- and 18-month-olds observed an agent expressing preference either matched (liking goldfish crackers and disliking broccoli) or unmatched (liking broccoli and disliking goldfish crackers) to their own preferences. When asked to offer the agent some food, younger children offered crackers (their own preference) regardless of condition; but older children offered what the agent preferred: crackers in the matched condition, and broccoli in the unmatched condition. Lucas et al. (2014) captured this developmental difference by proposing a shift in children's model of preference understanding. A simpler model assumes that all people have the same preferences, and the more complex model assumes that each individual has a distinct set of preferences. Lucas et al. (2014) ran simulations that captured how the accumulation of data would lead to a shift from the simpler to the more complex models. The simulations assumed that a child gradually observes more and more choices of her own, her parent and her sibling. In the beginning, her own preferences are broadly similar to her parents and siblings. The simpler and the more complex model capture the data equally well, therefore the simpler model is preferred based on Bayesian Occam's razor. As the child's observations grow, the simpler model fails to account for the differences in individuals' preferences, and the more complex model is now more probable. The simulations captured the experimental results in Repacholi and Gopnik (1997): younger children should be more likely to offer crackers to an agent with preferences unmatched to their own preferences, and older children should be more likely to offer broccoli. Consistent with the simulation results, a training study showed that after observing two experimenters expressing different preferences repeatedly, even 14-month-olds can represent others' preferences that are different from theirs (Doan, Denison, Lucas, & Gopnik, 2015). In addition, the model predicts that children who are in the process of shifting from the simpler to the more complex model should be sensitive to the strength of evidence indicating that an agent has a preference different from their own. This prediction is corroborated by evidence from 16-month-olds (Ma & Xu, 2011): when infants saw an agent choose six boring toys from a jar containing 13% of boring toys and 87% of interesting toys (strong evidence that the agent preferred the boring toy), they were more likely to offer a boring toy to the agent, compared to infants who saw an agent choose six boring toys from a jar containing 100% boring toys.

Inverse decision-making models have charted the developmental trajectory of infants' intuitive theories of preference. While 14-month-olds' performance in Repacholi and Gopnik (1997) appears irrational, the modeling work showed that it is because they have a simpler model of preference understanding, which is the most reasonable model given the data that they have observed. The modeling work further showed that infants can construct a new, more complex model of preference understanding through

Bayesian inductive learning. The transition from the simpler to the more complex model is predicted by both the amount and the strength of additional data that infants observe.

2.3 Naïve utility calculus

The last category of intuitive psychology models, the naïve utility calculus (NUC; Jara-Ettinger, Gweon, Schulz, & Tenenbaum, 2016; Jara-Ettinger, Schulz, & Tenenbaum, 2020), builds on both the inverse decision-making models and the inverse planning models. The NUC extended the inverse decision-making models by allowing joint inferences over costs and rewards (instead of just rewards). It extended the inverse planning models by adding a few additional levels of analysis in agents' action planning process, distinguishing between agents' desires, goals, intentions, and actions.

The general principle behind the NUC is that we reason about agents' actions based on the assumption that agents are intuitive utility maximizers. We assume that agents would maximize their rewards and minimize their costs when they choose their goals and actions. The model consists of a generative model that produces utility-maximizing behaviors given an agent's costs and rewards, and a mechanism that uses the generative model to infer an agent's costs and rewards from observed action sequences.

We will use a simple scenario shown in Fig. 3 to illustrate the model details. In this scenario, an astronaut is on an alien planet. She starts at the middle-left location on the map, and her goal is to get to the space station at the middle-right location. She can collect two types of care packages on

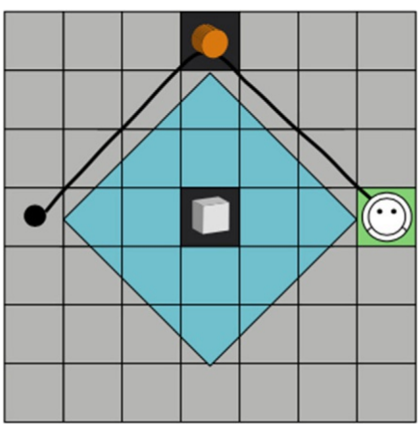


Fig. 3 A schematic of the astronaut scenario in Jara-Ettinger et al. (2020).

Fig. 4 The generative model of the naïve utility calculus (Jara-Ettinger et al., 2020).

the way, the orange one at the top and the white one in the middle. There are two different types of terrain, gray terrain and blue terrain.

The generative model is a hierarchical representation of how an agent's actions are produced (Fig. 4). An agent's desires are rewards associated with having different objects or helping different agents. In the astronaut scenario, the astronaut has different rewards for the space station, the yellow package, and the white package. These rewards, combined with beliefs about the location of the objects or agents determine the space of possible goals. Goals are defined as states of the world that the agent finds rewarding. In the astronaut scenario, the goals include getting to the space station, collecting the white package, and collecting the orange package. Next, the space of goals determines the space of intentions, which are ordered sequences of goals. The model only considers intentions that satisfy certain context-specific constraints (e.g., always have a specific goal as the final goal). In the astronaut scenario, the intentions must have the space station as the final goal. Thus, the space of intentions includes: (1) go to the space station; (2) collect the orange package, and go to the space station; (3) collect the white package, and go to the space station; (4) collect the orange package, collect the white package, and go to the space station; (5) collect the white package, collect the orange package, and go to the space station.

For each goal, the model computes an action policy that probabilistically maximizes utility; the action policy determines what actions the agent should take to achieve the goal state most efficiently. The action policy is computed through individual special-purpose Markov Decision Processes

(MDPs). The individual MDP uses the reward associated with the goal and plans movement in space. In the astronaut scenario, there are 49 positions on the map. At any time, the agent can move in one of eight directions (four cardinal directions and four diagonal directions). The cost of each action is determined by the type of terrain and length of the movement. A reward is obtained if the goal state is achieved (e.g., if the astronaut reached the orange package). The MDP computes an action policy that maps the goal states to actions with the highest utility, so that when the action policy is executed, the agent reaches the goal state as efficiently as possible. The utility of each intention is the sum of the goal's rewards minus the costs that the agent would incur to complete these goals in the specific order. The model selects an intention with the highest utility. Then, the selected intention is transformed into actions by executing each goal's action policy.

Another component of the model is a mechanism that uses Bayesian inference over the generative model to infer costs and rewards given a sequence of actions. The posterior probability of the cost and the reward functions is proportional to the product of the prior probability of the cost and the reward functions and the likelihood of observing the action sequence given the cost and the reward functions:

$$p(C, R|A) \propto l(A|C, R) \times p(C, R)$$

Since the terrains and the objects are novel, the model uses a uniform prior distribution for the cost function of each terrain, and a uniform prior distribution for the reward function of each object. The likelihood term is the probability that the generative model would produce the observed action sequence given each intention, summed over all possible intentions considered by the model:

$$l(A|C, R) = \sum_{I \in Intentions} p(A|I)p(I|C, R)$$

The likelihood term can be approximate with one or few high-probability intentions, rather than a sum over all possible intentions. The model approximates this Bayesian inference through Monte Carlo likelihood weighting. Instead of computing the posterior probability of all cost and rewards functions, the model samples n cost and reward functions from the prior distribution, and computes the likelihood of observing the action sequence under each sample of cost and reward functions.

The predictions of the NUC have been tested in adults with a variety of tasks. In one experiment (Jara-Ettinger et al., 2020), adults observed astronauts traveling to space stations in different environments (similar to Fig. 3). Then, they were asked to judge the agents' abilities to travel through each type of terrain, and their desires to collect each care package. The NUC predicted participants' judgments with quantitative accuracy, and showed more sensitivity than the alternative models, such as a heuristic model that infers cost from the amount of time an agent spends in each terrain and infers reward by the objects she collects. Other experiments (Jara-Ettinger et al., 2020) showed that the NUC also captured adults' performance when participants were asked to predict agents' future behaviors, to infer agents' knowledgeability about costs and rewards, and to reason about agents' social goals (helping other agents).

Some of these predictions have also been tested in children and infants. One set of studies (Jara-Ettinger, Gweon, Tenenbaum, & Schulz, 2015) found that 5- to 6-year-olds expected agents to maximize their utilities. They were able to use agents' choices under different situations to infer subjective rewards, and use agents' choices and subjective rewards to infer subjective costs. For instance, 5- to 6-year-olds learned that puppet A liked cookies over crackers, puppet B liked both equally, and that only one puppet could climb the tall box. When both puppets picked the cracker on the short box over the cookie on the tall box, children inferred that puppet A could not climb the tall box.

Another study (Jara-Ettinger, Tenenbaum, & Schulz, 2015) showed that even toddlers understand that different agents need to incur different costs for the same action. Seventeen- to 28-month-old toddlers observed a competent puppet make a toy play music on the first attempt and an incompetent puppet make the toy play music after many attempts. They preferred to play with the competent puppet, but when both puppets refused to help others activate the toy, they inferred that the incompetent puppet was nicer. Thus, consistent with the expectation that agents maximize utility, toddlers understand that failure to engage in a low-cost prosocial action implies that the agent has a lower reward associated with prosocial behaviors.

Infants understand that agents prefer goals that they achieved through costlier actions. In Liu et al. (2017), infants observed agents who were willing to incur different levels of cost to reach other agents. In one scenario (adapted from Gergely et al., 1995), the main character jumped over barriers of different heights to reach other agents. The main character was willing to jump over a low barrier, but not a medium barrier, to reach the blue agent; it was willing to jump over a medium barrier, but not a high barrier, to reach the yellow agent. Infants inferred that the main character preferred the yellow agent over the blue agent. Similar results were obtained when the cost was defined as the width of the gaps that the agent had to jump over, or the incline angle of ramps.

Children also understand that agents maximize their *expected* utilities, instead of *actual* utilities, since agents might be ignorant about their own costs and rewards (Jara-Ettinger, Floyd, Tenenbaum, & Schulz, 2017). Four- to five-year-olds understand that knowledgeable agents are more likely to choose high utility options and their choices are more likely to be stable, compared to agents who are ignorant about their own rewards or costs. For instance, children were presented with two puppets and two novel fruits. They were told that one puppet had tasted both fruits before and the other had not. Both puppets chose the same fruit; one said "Yum" and the other said "Yuck." Children inferred that the puppet who said "Yuck" had not tasted the fruits before.

The NUC model has been extended to capture the reasoning behind teaching decisions (Bridgers, Jara-Ettinger, & Gweon, 2020). Children learned about two toys that varied in discovery rewards and discovery costs, and the experimenter (the teacher) could only teach one of the toys to a naïve learner. Children helped the experimenter decide which toy she should teach the learner, and which toy she should let the learner discover on her own. The model makes this decision by maximizing the learner's expected utility of learning from instruction and exploration. Thus, the learner's utility is the activation reward of the instructed toy and the discovery reward of the explored toy, subtracted by the activation cost of the instructed toy and the discovery cost of the explored toy. Five- to seven-year-olds' behaviors were consistent with the model's predictions—they were more likely to teach a toy with a higher reward than a toy with a lower reward when the costs were equal; they were more likely to teach a toy with a higher cost than a toy with a lower cost when the rewards were equal; when one toy had a higher reward and a lower cost and the other toy had a lower reward and a higher cost, children were increasingly more likely to teach the higher cost toy as the difference in costs increased.

Lastly, Meng and Xu (2020, 2021) have shown that the rational inference behind NUC alone can "reproduce" observed disparity among unbiased people. For example, when people observe police officers stopping and searching Black people at higher rates compared to the other groups, they may infer that Black people are more likely to commit crimes because they assume that the police officers would maximize their utilities by searching groups with higher crime rates. In one study (Meng & Xu, 2021), adults observed a knowledgeable agent, a border patrol officer on a planet, using a costly method (a scanner) to check whether aliens from other planets had stolen gems. The officer knew the "theft rate" of aliens from each planet and checked aliens from a series of planets at different rates. In one condition, participants also observed the sample hit rates—the number of aliens who actually stole gems among the ones the officer checked. In the critical trials, the check rate and the hit rate conflicted (e.g., the hit rate was low when the check rate was high). Then, participants were asked to infer the theft rate of all aliens on each planet. Adults' inferences were consistent with the predictions of the NUC models. When they were only given check rates, they inferred higher theft rates as the officer checked the aliens more often. When they were given both check rates and hit rates, participants rationally integrated the two types of information in their inferences (e.g., they inferred moderate theft rate for groups that were checked often but had a low sample hit rate). Thus, observing a knowledgeable, utility-maximizing agent checking different groups at different rates leads to the reproduction of disparities. Providing sample hit rates might reduce this negative consequence.

The NUC provides a unified account for reasoning about agents' actions across a variety of contexts. We assume that others choose actions that maximize their utilities—they tradeoff between costs and rewards in a precise manner. Furthermore, this assumption is consistent across development. Even infants use this assumption to reason about agents' actions in simple scenarios. The NUC also motivated many additional behavioral experiments that tested different aspects of this unified theory.

A future direction for both the NUC models and the inverse planning models is to incorporate mental costs of planning in the models. One reason that agents do not always act rationally is that planning complex sequences of actions takes time and mental resources. When people do not have enough time and resources to engage in planning, they execute non-optimal plans and inefficient actions. People might also tradeoff between the mental cost of planning and the reward of finding an optimal plan. That is, they will only engage in planning if the optimal plan leads to an increase in rewards that offsets the mental costs of planning. There are two ways to extend the NUC and inverse planning models in this direction. First, the models can move toward the algorithmic level of analysis, and consider the plausible algorithms that are underlying human reasoning given the limited amount of cognitive resources. For instance, Lieder and Griffiths (2020) proposed a new modeling paradigm called the *resource-rational analysis*.

The *recourse-rational analysis* assumes that people make rational use of their limited cognitive resources, maximizing the utility of the decision while minimizing the costs of the decision-making process. Second, a variable representing the mental costs of planning can be directly added to the models. The models can tradeoff between the mental costs and the rewards of finding an optimal plan to determine the amount of planning the agent should engage in.

In sum, the three categories of models reviewed above have successfully captured infants' and children's abilities to make mental state inferences across several key phenomena in the domain of intuitive psychology. These models showed that infants, children, and adults reason about agents based on intuitive theories of psychology that specify how agents plan their actions. Utility calculation—the assumption that agents choose actions that maximize their rewards and minimize their costs—is a critical component in inverse planning and NUC models, suggesting that it might be a unifying principle underlying our intuitive theory of psychology.

3. Pedagogical reasoning and epistemic trust

The social world also provides children with the opportunity to learn from others. Teaching allows quick transmission of important knowledge. However, choosing the best examples to teach and learning the correct hypotheses based on the examples are nontrivial inferential problems. What kinds of mechanisms underlie our abilities to learn from other people and teach other people? This question is addressed by the pedagogical model. In addition, not everyone is the best teacher to learn from. How do we decide whom we should trust and learn from? This question is addressed by the epistemic trust model.

3.1 Pedagogical model

When children learn about the world, they need to infer the correct hypotheses based on the enormous amount of data that they observe in the world. Pedagogical situations have an important impact on this process, because in these situations, data are not generated randomly; instead, the teacher is choosing data to transmit to the learner for the purpose of teaching a particular hypothesis (Csibra & Gergely, 2009). How can a teacher optimally teach a hypothesis to a learner? Imagine you want to teach a child the word "dog." You can use the word "dog" to refer to three golden retrievers, or use the word "dog" to refer to a golden retriever, a dalmatian, and a border collie. Intuitively, the second set of examples is better than the first one. The child is more likely to infer the correct hypothesis, that "dog" refers to the basic level category *dog*, given the second set of examples. Why is that the case? When you choose three golden retrievers as examples, the child would be more likely to believe in the hypothesis that "dog" refers to the subordinate level category *golden retriever*, rather than the hypothesis that it refers to the basic level category *dog*. If the latter hypothesis were true, you would have chosen examples of other kinds of dogs as well. However, the child will only engage in this kind of reasoning if she assumes that you are intentionally choosing the examples to help her learn. In contrast, if the child overheard you say "dog" three times while pointing to three golden retrievers, the child would be less certain that "dog" refers to *golden retriever*. Because the examples are generated randomly, both hypotheses ("dog" refers to *golden retriever*, and "dog" refers to *dog*) are consistent with the data she observed. Therefore, in pedagogical communications, the learner needs to reason about the teacher's intentions, and distinguish between when the examples are provided for the purpose of helping her learn and when the examples are generated randomly. The teacher also needs to reason about what the learner would infer given different examples, and choose the examples that are most helpful for the learner.

The natural pedagogy perspective (Csibra & Gergely, 2009) proposes that such pedagogical communications are achieved by a specifically adapted human communication system called "natural pedagogy." "Natural pedagogy" takes place when communications are accompanied by ostensive cues such as eye contact, pointing, and child-directed speech. Most importantly, this perspective proposes that the knowledge being transmitted in such contexts is kind-relevant and generalizable. Indeed, evidence suggests that both infants and preschoolers expect to learn generalizable knowledge when communications are accompanied by ostensive cues (e.g., Butler & Markman, 2012; Egyed, Király, & Gergely, 2013).

In contrast, the pedagogical model proposed by Shafto, Goodman, and Griffiths (2014) does not make any assumptions about the kinds of knowledge being transmitted in pedagogical communications. Instead, it focuses on the learner's and the teacher's ability to reason about the mental states of each other. In particular, the learner reasons about the process by which the teacher chooses the data as she updates her beliefs; the teacher reasons about the learner's belief updating process and chooses data that will be most helpful to the learner.

How should the teacher sample data that are helpful to the learner? The literature on concept learning has distinguished between weak sampling and strong sampling. In weak sampling, data are randomly selected from all possible examples, and are labeled as to whether they are true of the target hypothesis (e.g., Hsu & Griffiths, 2009). In strong sampling, data are randomly selected from the set of examples that are true of the hypothesis (e.g., Xu & Tenenbaum, 2007a, 2007b). However, neither of these two types of sampling could capture the sampling process that underlies teaching. A key assumption of the pedagogical model is that the teacher is engaging in *pedagogical sampling*—choosing data that will maximize the learner's belief in the correct hypothesis, that is, the posterior probability of the correct hypothesis. The model further assumes that the learner knows that data are sampled by a helpful teacher, and rationally updates her belief.

Based on these two assumptions, the pedagogical model formalizes a system of equations, specifying the distribution from which the teacher generates data (the sampling distribution), and the process by which the learner updates her belief given data. Critically, these equations depend on each other—the teacher's sampling distribution depends on how the learner will update her belief given data; the learner's belief updating process depends on the teacher's sampling distribution. The model solves this system of equations using a mathematical method called fixed-point iteration. This method is analogous to a process of recursive mental state reasoning (although the actual psychological mechanisms used to solve this problem do not necessarily involve explicit recursion). When the learner updates her belief given data, she needs an estimation of the likelihood that the teacher generated the data given the true hypothesis; to do this, she has to make an assumption about the teacher's sampling distribution (e.g., weak sampling, strong sampling, or pedagogical sampling); if pedagogical sampling is assumed, this process depends on the teacher's assumption about how the learner will update her belief; if the teacher assumes the learner will rationally update her belief, this assumption, in turn, depends on the learner's assumption about the teacher's sampling distribution, and so on. This recursive reasoning will eventually converge, at which point we have the solution to the system of equations. The model achieves this solution by first specifying an initial distribution from which the teacher generates data—the initial sampling distribution assumes unbiased random sampling. Depending on whether negative evidence is possible, it will be either weak sampling (random sampling from all possible examples) or strong sampling (random sampling from all examples true of the hypothesis). Then, the fixed-point iteration process described above will transform this initial distribution into a solution satisfying all model assumptions.

For instance, in one task in Shafto et al. (2014), a teacher teaches a rule-based concept (a rectangle on a board) to a learner. In one condition, the teacher can provide two positive examples (i.e., two points inside the rectangle) to help a learner infer the correct rectangle. The model assumes that the teacher's initial sampling distribution is strong sampling—the teacher is equally likely to choose any points inside the rectangle. Then, the model makes predictions about the hypotheses that the learner will infer given the assumption that the teacher samples from this distribution. The predictions would show that some examples are more likely than others to allow the learner to infer the correct hypothesis (e.g., points near the corners of the rectangle will be more likely to lead to the correct hypothesis compared to other examples). Then, the model updates the teacher's sampling distribution by increasing the probability of sampling from examples that are more likely to lead to the correct hypothesis (e.g., given the new distribution, the teacher will be more likely to choose points closer to the corners). Then, the model repeats the processes of predicting the hypotheses that the learner will infer, and updating the teacher's sampling distribution based on the predictions. Eventually, the teacher's sampling distribution will converge to examples that are most likely to lead to the correct hypothesis (e.g., pairs of points at two opposite corners). Indeed, adults who played the role of the teacher in the experiments chose these examples to teach the learner. Other participants played the role of the learner. When learners were told that the examples were generated by a helpful teacher, the rectangles that they inferred were more likely to have the positive examples at the corners. That is, they understood the process by which the teacher chose the examples, and were more likely to infer the correct rectangle. When learners thought the examples were generated by the computer, the rectangles they inferred did not show this specific pattern, and they were less likely to infer the correct rectangle. Shafto et al. (2014) examined the model predictions in two other tasks on prototype concepts and causally structured concepts, and showed that adults' teaching and learning in these tasks were well captured by the pedagogical model. Furthermore, this model can be applied to teaching and learning any concept as long as the hypothesis space can be specified.

The general prediction of the pedagogical model is that a teacher should provide examples that would maximize the learner's belief in the correct hypothesis. A few specific predictions follow from this general prediction:

first, the examples that a teacher provided should be exhaustive, and the learner can infer that any hypotheses for which the teacher did not provide evidence are not true; second, a teacher would only provide necessary examples, and the true hypothesis should be consistent with all examples, rather than a portion of the examples; third, when there are infinite numbers of possible examples for a hypothesis, such as points in the rectangle in the task in Shafto et al. (2014), the most diverse set of examples should be preferred (e.g., points on the opposite corners of the rectangle). Studies with young children have supported these model predictions. First, Bonawitz et al. (2011) found that 4- to 5-year-olds expected a teacher's examples to be exhaustive. When a knowledgeable teacher demonstrated one function on a toy, they explored the toy less later, compared to when the function was demonstrated by an ignorant teacher. Children inferred that a knowledgeable teacher would demonstrate all functions on the toy; since she only demonstrated one, children did not expect to find additional functions through exploration. Second, Buchsbaum, Gopnik, Griffiths, and Shafto (2011) found that 3- to 5-year-olds expected a teacher to only provide necessary examples. In this study, children learned about a causally structured concept—the sequence of actions that leads to an effect. The teacher was either knowledgeable about the toy and taught the child how it worked by demonstrating several three-action sequences (pedagogical condition), or the teacher was ignorant about the toy and demonstrated the same three-action sequences while she tried to figure out how it works (non-pedagogical condition). In some trials, the statistical evidence indicated that only two of the actions in the sequences were the overlapping cause of the effect. In accordance with the statistical evidence, children in the non-pedagogical condition produced the two-action sequence. However, children in the pedagogical condition were more likely to over-imitate and produce the three-action sequence. That is, children inferred that all three actions were necessary, because they assumed that the teacher would not demonstrate superfluous actions if she was helping them learn. Last, Rhodes, Gelman, and Brickman (2010) found that 6-year-olds prefer to teach with a diverse set of examples. Participants taught another child a novel property that was true of a subset of animals (e.g., dogs). They found that children were more likely to choose a diverse set of examples (e.g., a golden retriever, a dalmatian, and a collie) than a non-diverse set of examples (e.g., three dalmatians).

The development of the pedagogical model has also led to new behavioral investigations. For instance, given the important role that mental state reasoning plays in the pedagogical model, Bass et al. (2019) examined the link between preschoolers' ability to select evidence to correct others' false beliefs and their theory of mind abilities. Three- to four-year-olds first learned the kind of blocks that activate a toy (e.g., red blocks activate the toy). Then, they provided evidence in the presence of a confederate with a false belief (e.g., the confederate believed square blocks activated the toy). Children with better evidence selection ability were more likely to select evidence contradicting the confederate's false belief (e.g., showing that a red, circle block activated the toy). They found that there is a correlation between children's evidence selection ability and ToM ability, above and beyond effects of age and other cognitive abilities. Furthermore, a 6-week training of pedagogical evidence selection improved children's ToM ability. In another study, Gweon, Shafto, and Schulz (2018) investigated the concept of informativeness in teaching in 5- to 6-year-olds. They found that children preferred teachers who provided a sufficient number of demonstrations based on the learners' knowledge levels, and modulated their own teaching in similar fashions. The pedagogical model assumes that the teacher should maximize the probability that the learner believes in the correct hypothesis, that is, they should always generate more data until no additional data could benefit the learner. However, even children understand that teachers should not be overinformative. Gweon et al. (2018) extended the pedagogical sampling assumption by incorporating the cost of information transmission—a rational teacher should maximize the utility of information, instead of the reward of information.

Next, we will identify two sets of behavioral studies in cognitive development that could be integrated with the pedagogical model. The first set of studies supported the natural pedagogy assumption, that knowledge transmitted through pedagogical communication is generalizable (e.g., Butler & Markman, 2012; Egyed et al., 2013). As Shafto et al. (2014) also discussed, their model could be extended to incorporate the natural pedagogy assumption. For instance, the model could specify that in pedagogical situations, the learner has stronger prior beliefs for hypotheses about generalizable concepts than hypotheses about nongeneralizable concepts. The teacher would assume that the learner has these prior beliefs when she engages in pedagogical sampling.

The second set of studies examined children's normative inferences about intentional actions. Schmidt, Butler, Heinz, and Tomasello (2016) found that 3-year-olds are "promiscuous normativitists," that is, they inferred social norms from single, spontaneous human actions, in the absence of any linguistic or behavioral cues indicating that the actions might be generic or normative. Furthermore, Butler, Schmidt, Bürgel, and Tomasello (2015) found that 3-year-olds' normative inferences were even stronger when the actions were demonstrated pedagogically, compared to when they were demonstrated intentionally. These findings can be analyzed in the pedagogical model framework: Children assign stronger prior probabilities to normative hypotheses compared to non-normative hypotheses about any intentional actions—they believe that intentionally demonstrated actions are more likely to be normative. Pedagogical demonstration of an action further increases the probability that the normative hypothesis is correct. Since the learner assumes that the teacher knows their prior beliefs (e.g., that an intentionally demonstrated action is more likely to be normative), the fact that the teacher chose to demonstrate this particular action increases the learner's posterior belief that this action should be normative.

Lastly, we will point out one limitation of the pedagogical model. The model assumes the teacher always has accurate representations of the learner's hypothesis space, which might not be true in naturalistic teaching situations. For instance, Aboody, Velez-Ginorio, Santos, and Jara-Ettinger (2018) showed that the teacher's representation of the learner's hypothesis space might be simpler than the actual hypothesis space considered by the learner. They designed a task in which a participant taught another participant the activation rule for a toy. Given the evidence that the "teachers" provided, only about half of the "learners" inferred the correct rule. Aboody and colleagues entered the evidence provided by the "teacher" participants into a computational model similar to the pedagogical model and examined the model's inference of the activation rule. The accuracy of the model inference was much higher (75%) when the model learned under a simple hypothesis space (e.g., only allowing rules containing single blocks or two blocks), than under a complex hypothesis space (25%). Furthermore, constraining participants' hypothesis space led to more accurate learning given the same evidence provided by the "teachers."

In sum, the pedagogical model describes how teachers choose data and how learners update their beliefs in pedagogical situations. Past studies on adults' and children's learning and teaching are consistent with the model predictions. The model has led to interesting behavioral investigations about the relationship between ToM and pedagogical evidence selection, as well as children's understanding of informativeness in teaching. Future work could integrate the natural pedagogy perspective with the pedagogical model, and use the pedagogical model to capture children's normative inferences about intentional actions. The model can also be extended to capture teaching and learning when the teacher does not have an accurate representation of the learner's hypothesis space.

3.2 Epistemic trust model

While the pedagogical model only focuses on learning in situations where the teacher is always knowledgeable and helpful, the epistemic trust model (Shafto, Eaves, Navarro, & Perfors, 2012) captures the phenomenon of learning from others in a broad range of situations. When children do not know who is knowledgeable and helpful, how do they decide whom to trust and learn about the world simultaneously? The epistemic trust model developed by Shafto et al. (2012) focuses on a particular task that has been studied extensively in the epistemic trust literature—learning a new word label for a novel object from informants.

The epistemic trust model includes a generative model about how an informant chooses a label to provide; a learner observes the provided label, and simultaneously infers the true state of the world (i.e., the actual label for the object) and the informant's knowledgeability and helpfulness. The generative model specifies that the informant chooses a label based on her belief and helpfulness. Her belief depends on her knowledgeability and the true state of the world. The model assumes that the knowledgeability variable is binary: A knowledgeable informant always believes in the actual label, and a non-knowledgeable informant is equally likely to believe in any label in a set of possible labels. The helpfulness variable is also assumed to be binary: A helpful informant selects the label that maximizes the probability that the learner forms the same belief as the informant, whereas an unhelpful informant minimizes this probability.

Shafto et al. (2012) compared the fits of two models with children's performance in past studies. The Knowledge & Intent model includes parameters (free parameters that were fitted to data) that reflect children's prior beliefs about informants' knowledgeability and helpfulness on average, and the variability of knowledgeability and helpfulness across informants. The Knowledge-only model also included the knowledgeability parameters (fitted to data), and the helpfulness parameters were fixed to reflect the assumption that informants are always helpful.

Past studies have demonstrated 4-year-olds' competence in various epistemic trust tasks. However, 3-year-olds' performance in some of these tasks differed from 4-year-olds. Different studies suggest slightly different interpretations for this developmental change. In Pasquini, Corriveau, Koenig, and Harris (2007), children were asked to endorse novel labels provided by two informants with different previous accuracies on labeling familiar objects (100% vs 0%; 100% vs 25%; 75% vs 0%; 75% vs 25%). While 3-year-olds trusted the more accurate informants only when one of the informants was always accurate, 4-year-olds trusted the more accurate informants in all conditions. These results suggest that 3-year-olds might be less sensitive to the relative frequency of errors compared to 4-year-olds. In Corriveau and Harris (2009), children were presented with familiar and unfamiliar informants, and the informants provided correct or incorrect labels for familiar objects in the experiments. Although both 3- and 4-year-olds initially trusted familiar over unfamiliar informants, when the familiar informant provided incorrect labels for objects and the unfamiliar informant provided correct labels, only 4-year-olds trusted the unfamiliar informant more. These results suggest that younger children might have trouble integrating new accuracy information into judgments about familiar informants. In Mascaro and Sperber (2009), 4-year-olds decided not to trust an informant whom they were explicitly told to be a big liar, whereas 3-year-olds continue to trust that informant. This result suggests that younger children might ignore information about informants' intent. Lastly, in Corriveau, Fusaro, and Harris (2009), both 3- and 4-year-olds were more likely to trust an informant who previously referred to a novel object with a label that agreed with the majority of the group, compared to an informant who previously disagreed with the majority of the group.

Across all 4 studies, 4-year-olds' behaviors were best captured by the Knowledge & Intent model. The best-fitting parameters revealed that 4-year-olds believe that informants are knowledgeable and helpful on average, but different informants vary in their knowledgeability and helpfulness. However, across all four studies, 3-year-olds' behaviors were captured equally well by the Knowledge & Intent model and the Knowledge-only model. Moreover, the best-fitting parameters of the Knowledge & Intent model for 3-year-olds were consistent with an assumption that people are uniformly helpful. In other words, 3-year-olds believe that all informants are helpful, but some informants are more knowledgeable than others. Thus, the modeling results provided a unifying alternative explanation for the developmental changes observed in 3- and 4-year-olds' performance in various tasks—that 3- and 4-year-olds differ in their assumptions about the helpfulness of informants. Future work could examine whether this developmental shift can be captured by Bayesian inductive learning, that is, whether the shift to the new assumption reflects an integration of children's prior beliefs (i.e., all informants are helpful) and the new data they observe (e.g., some informants are not helpful).

So far, the epistemic trust model has only been used to capture children's behaviors in a particular task: when children were given information about the reliability of informants, they selectively learn novel labels from reliable informants. However, children's epistemic reasoning abilities have been demonstrated in various tasks. Next, we identify three lines of behavioral studies that could be integrated with the epistemic trust model.

First, Schütte, Mani, and Behne (2020) have demonstrated 5-year-olds' ability to use informants' reliability to make epistemic trust judgments retrospectively. In their study, children first observed conflicting testimonies provided by two unfamiliar informants—they used the same novel label to refer to different novel objects. Then, they received new information about the informants' reliability: one informant consistently labeled familiar objects accurately, and the other consistently labeled them inaccurately. Children retrospectively inferred that the testimony (i.e., the referent of the novel label) provided by the reliable informant was more likely to be correct. The epistemic trust model could be used to capture these results. At first, children had no information about the knowledgeability and helpfulness of the two informants, and they might infer that the two conflicting testimonies were equally likely to be true. When they received new information about the informant's knowledgeability, their original inferences were updated—now they inferred that the knowledgeable informant was more likely to have provided the correct testimony.

Another study by Liberman and Shaw (2020) found that children understand that people can be biased in their testimony about friends or enemies. When 3- to 13-year-olds heard a negative testimony about a person (e.g., "she is bad at soccer"), they judged that the target person was worse at soccer if the testimony was provided by a friend, compared to an enemy. When they heard a positive testimony about a person, they made more positive judgments about the target person's ability if the testimony was provided by an enemy, as opposed to a friend. These results could be captured by the epistemic trust model with a slight modification. The helpfulness variable in the original model could be changed to a variable about general intention. The nature of the social relationship between the informant and the target of the testimony affects the informant's intention (e.g., the informant could be positively or negatively biased). Observers can infer the true state of the world (e.g., the target's true ability) based on the testimony and the social relationship between the informant and the target (i.e., they can correct for the informant's bias).

Lastly, studies by Butler, Schmidt, Tavassolie, and Gibbs (2018) and Butler, Gibbs, and Tavassolie (2020) showed that children reason about the process by which claims are made, above and beyond reasoning about informants' knowledgeability and helpfulness. Children as young as 3 years of age judged verified claims (e.g., when a person looked inside a container and made a claim about what is inside the container) to be more acceptable than claims that have not been verified (e.g., when the person chose not to look inside the container and made the same claim) (Butler et al., 2018). In addition, 6- and 7-year-olds, but not younger children, treated verification as more important than an informant's past history of accuracy when they judged whether a claim should be trusted (Butler et al., 2020). They were more likely to trust a verified claim provided by an informant who previously provided inaccurate word labels, compared to an unverified claim provided by an informant who previously provided accurate word labels. These results could help extend the existing epistemic trust model. For instance, an additional variable, the method that informants use to form their beliefs, could be incorporated into the model. This variable would determine whether or not the informant's belief would match the true state of the world. For instance, if a verification process was used, the informant's belief is more likely to match the true world state, and therefore the informant's testimony is more likely to be accurate. Modeling these results could also shed light on the developmental change in children's ability to integrate verification and past accuracy information in their epistemic trust judgments.

In sum, the epistemic trust model describes how children reason about the testimony provided by informants to learn about the world and to infer the reliability of informants. Past studies on children's ability to learn novel labels from informants are well captured by this model. Recent studies have demonstrated children's abilities to retrospectively update their beliefs about testimonies, to reason about biased testimony, and to reason about the process by which testimonies are made. These new findings are consistent with the predictions of the epistemic trust model, and suggest ways that the model could be extended to reflect children's sophisticated reasoning in epistemic judgments.

4. Conclusion

In the first half of the chapter, we reviewed a growing body of Bayesian probabilistic models on reasoning about agents' mental states and actions. Comparisons of these models with human performances have shown that infants, children, and adults reason about agents based on intuitive theories of psychology that specify how agents plan their actions. Examining the transition of infants' theories of intuitive preference has shown that infants can construct new concepts and theories through Bayesian inductive learning, consistent with the rational constructivism framework.

In the second half of the chapter, we reviewed Bayesian probabilistic models on pedagogical reasoning and epistemic trust. The pedagogical model has shown that teachers and learners engage in optimal teaching and learning by reasoning about the mental states of each other. The teacher provides evidence that would maximize the learner's belief in the correct hypothesis. The epistemic trust model has shown that children selectively trust informants by inferring the knowledgeability and helpfulness of informants. The model also revealed a shift between 3- and 4-year-olds' assumptions of the helpfulness of informants, and whether this shift can be captured through Bayesian inductive learning within the rational constructivism framework remains to be explored.

Taking these two bodies of work together, we found that the abilities to learn about others—inferring others' mental states—emerge during infancy and guide our reasoning throughout childhood and adulthood, but the abilities to learn from others—pedagogical reasoning and epistemic reasoning—emerge later during preschool. How should we reconcile the early competence in mental states reasoning and the later developing abilities in pedagogical and epistemic reasoning? As a concrete example, infants as young as 10-month-olds are capable of socially evaluating prosocial and antisocial agents based on mental state inferences, but 3-year-olds assume that informants are uniformly helpful when they learn new words from them. Future work should further integrate these two lines of research. For instance, components of the intuitive psychology models could be incorporated into the pedagogical and epistemic trust models to understand how mental state reasoning relates to and potentially contributes to the development of pedagogical and epistemic reasoning.

Given the scope of this chapter, we could not elaborate on many other social learning models that might be of interest to researchers in child development. But before we close, we would like to highlight one other category of models that focus on learning and reasoning about social groups, namely the social-structure learning models (Gershman & Cikara, 2020; Gershman, Pouncy, & Gweon, 2017; Lau, Pouncy, Gershman, & Cikara, 2018; Martinez, Feldman, Feldman, & Cikara, 2021). These models sort individuals into latent groups based on the assumption that individuals in the same group tend to behave similarly. Social-structure learning models provide a promising approach to investigate many issues on children's reasoning about social groups, such as intergroup bias, stereotyping, essentialist beliefs, and so on.

Bayesian probabilistic models have provided formal accounts of many aspects of cognitive, language, and social-cognitive development, and inspired new empirical investigations. We hope that this review will inspire fruitful future research to tackle the open questions in both child development and computational modeling.

References

Aboody, R., Velez-Ginorio, J., Santos, L., & Jara-Ettinger, J. (2018). In T. T. Rogers, M. Rau, X. Zhu, & C. W. Kalish (Eds.), *When teaching breaks down: Teachers rationally select what information to share, but misrepresent learners' hypothesis spaces Proceedings of the 40th annual conference of the cognitive science society* (pp. 70–75). Austin, TX: Cognitive Science Society.

Aguiar, A., & Baillargeon, R. (1999). 2.5-month-old infants' reasoning about when objects should and should not be occluded. *Cognitive Psychology*, *39*(2), 116–157. https://doi.org/10.1006/cogp.1999.0717.

Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, *1*(4), 0064. https://doi.org/10.1038/s41562-017-0064.

Bass, I., Gopnik, A., Hanson, M., Ramarajan, D., Shafto, P., Wellman, H., et al. (2019). Children's developing theory of mind and pedagogical evidence selection. *Developmental Psychology*, *55*(2), 286–302. https://doi.org/10.1037/dev0000642.

Blythe, P. W., Todd, P. M., & Miller, G. F. (1999). How motion reveals intention: Categorizing social interactions. In I. G. Gigerenzer, P. M. Todd, & the ABC research group (Eds.), *Simple heuristics that make us smart* (pp. 257–285). Oxford University Press.

Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition*, *120*(3), 322–330. https://doi.org/10.1016/j.cognition.2010.10.001.

Boyd, R., Richerson, P. J., & Henrich, J. (2011). The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(Suppl_2), 10918–10925. https://doi.org/10.1073/pnas.1100290108.

Bridgers, S., Jara-Ettinger, J., & Gweon, H. (2020). Young children consider the expected utility of others' learning to decide what to teach. *Nature Human Behaviour*, *4*, 144–152. https://doi.org/10.1038/s41562-019-0748-6.

Buchsbaum, D., Gopnik, A., Griffiths, T. L., & Shafto, P. (2011). Children's imitation of causal action sequences is influenced by statistical and pedagogical evidence. *Cognition*, *120*(3), 331–340. https://doi.org/10.1016/j.cognition.2010.12.001.

Butler, L. P., Gibbs, H. M., & Tavassolie, N. S. (2020). Children's developing understanding that even reliable sources need to verify their claims. *Cognitive Development*, *54*, 100871. https://doi.org/10.1016/j.cogdev.2020.100871.

Butler, L. P., & Markman, E. M. (2012). Preschoolers use intentional and pedagogical cues to guide inductive inferences and exploration: Pedagogical cues and exploration. *Child Development*, *83*(4), 1416–1428. https://doi.org/10.1111/j.1467-8624.2012.01775.x.

Butler, L. P., Schmidt, M. F. H., Bürgel, J., & Tomasello, M. (2015). Young children use pedagogical cues to modulate the strength of normative inferences. *British Journal of Developmental Psychology*, *33*(4), 476–488. https://doi.org/10.1111/bjdp.12108.

Butler, L. P., Schmidt, M. F. H., Tavassolie, N. S., & Gibbs, H. M. (2018). Children's evaluation of verified and unverified claims. *Journal of Experimental Child Psychology*, *176*, 73–83. https://doi.org/10.1016/j.jecp.2018.07.007.

Corriveau, K. H., Fusaro, M., & Harris, P. L. (2009). Going with the flow: Preschoolers prefer nondissenters as informants. *Psychological Science*, *20*(3), 372–377. https://doi.org/10.1111/j.1467-9280.2009.02291.x.

Corriveau, K., & Harris, P. L. (2009). Choosing your informant: Weighing familiarity and recent accuracy. *Developmental Science*, *12*(3), 426–437. https://doi.org/10.1111/j.1467-7687.2008.00792.x.

Csibra, G., Bı´ró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science*, *27*(1), 111–133. https://doi.org/10.1016/S0364-0213(02)00112-X.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 148–153. https://doi.org/10.1016/j.tics.2009.01.005.

Csibra, G., Gergely, G., Bı´ró, S., Koós, O., & Brockbank, M. (1999). Goal attribution without agency cues: The perception of 'pure reason' in infancy. *Cognition*, *72*(3), 237–267. https://doi.org/10.1016/S0010-0277(99)00039-6.

Doan, T., Denison, S., Lucas, C., & Gopnik, A. (2015). *Learning to reason about desires: An infant training study* In D. C. Noelle et al. (Eds.), Proceedings for the Annual Meeting of the Cognitice Science Society (pp. 578–583). Austin, TX: Cognitive Science Society.

Egyed, K., Király, I., & Gergely, G. (2013). Communicating shared knowledge in infancy. *Psychological Science*, *24*(7), 1348–1353. https://doi.org/10.1177/0956797612471952.

Gao, T., Newman, G. E., & Scholl, B. J. (2009). The psychophysics of chasing: A case study in the perception of animacy. *Cognitive Psychology*, *59*(2), 154–179. https://doi.org/10.1016/j.cogpsych.2009.03.001.

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational action. *Trends in Cognitive Sciences*, 7(7), 287–292. https://doi.org/10.1016/S1364-6613(03)00128-1.

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*(2), 165–193. https://doi.org/10.1016/0010-0277(95)00661-H.

Gershman, S. J., & Cikara, M. (2020). Social-structure learning. *Current Directions in Psychological Science*, *29*(5), 460–466. https://doi.org/10.1177/0963721420924481.

Gershman, S. J., Pouncy, H. T., & Gweon, H. (2017). Learning the structure of social influence. *Cognitive Science*, *41*, 545–575. https://doi.org/10.1111/cogs.12480.

Gopnik, A., & Wellman, H. M. (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychological Bulletin*, *138*(6), 1085–1108. https://doi.org/10.1037/a0028044.

Gweon, H., Shafto, P., & Schulz, L. (2018). Development of children's sensitivity to over-informativeness in learning and teaching. *Developmental Psychology*, *54*(11), 2113–2125. https://doi.org/10.1037/dev0000580.

Hamlin, J. K., Ullman, T., Tenenbaum, J., Goodman, N., & Baker, C. (2013). The mentalistic basis of core social cognition: Experiments in preverbal infants and a computational model. *Developmental Science*, *16*(2), 209–226. https://doi.org/10.1111/desc.12017.

Hamlin, J. K., & Wynn, K. (2011). Young infants prefer prosocial to antisocial others. *Cognitive Development*, *26*(1), 30–39. https://doi.org/10.1016/j.cogdev.2010.09.001.

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*(7169), 557–559. https://doi.org/10.1038/nature06288.

Hamlin, J. K., Wynn, K., Bloom, P., & Mahajan, N. (2011). How infants and toddlers react to antisocial others. *Proceedings of the National Academy of Sciences*, *108*(50), 19931–19936. https://doi.org/10.1073/pnas.1110306108.

Hsu, A., & Griffiths, T. (2009). Differential use of implicit negative evidence in generative and discriminative language learning. In Y. Bengio, D. Schuurmans, J. Lafferty, C. Williams, & A. Culotta (Eds.), *Advances in neural information processing systems* (Vol. 22). Curran Associates, Inc. https://proceedings.neurips.cc/paper/2009/file/0f96613235062963ccde717b18f97592-Paper.pdf.

Jara-Ettinger, J., Floyd, S., Tenenbaum, J. B., & Schulz, L. E. (2017). Children understand that agents maximize expected utilities. *Journal of Experimental Psychology: General*, *146*(11), 1574–1585. https://doi.org/10.1037/xge0000345.

Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends in Cognitive Sciences*, *20*(8), 589–604. https://doi.org/10.1016/j.tics.2016.05.011.

Jara-Ettinger, J., Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2015). Children's understanding of the costs and rewards underlying rational action. *Cognition*, *140*, 14–23. https://doi.org/10.1016/j.cognition.2015.03.006.

Jara-Ettinger, J., Schulz, L. E., & Tenenbaum, J. B. (2020). The naïve utility calculus as a unified, quantitative framework for action understanding. *Cognitive Psychology*, *123*, 101334. https://doi.org/10.1016/j.cogpsych.2020.101334.

Jara-Ettinger, J., Tenenbaum, J. B., & Schulz, L. E. (2015). Not so innocent: Toddlers' inferences about costs and culpability. *Psychological Science*, *26*(5), 633–640. https://doi.org/10.1177/0956797615572806.

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*(1), 1–19. https://doi.org/10.1016/0010-0277(91)90045-6.

Kushnir, T., Xu, F., & Wellman, H. M. (2010). Young children use statistical sampling to infer the preferences of other people. *Psychological Science*, *21*(8), 1134–1140. https://doi.org/10.1177/0956797610376652.

Lau, T., Pouncy, H. T., Gershman, S. J., & Cikara, M. (2018). Discovering social groups via latent structure learning. *Journal of Experimental Psychology: General*, *147*(12), 1881–1891. https://doi.org/10.1037/xge0000470.

Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition*, *25*(3), 265–288. https://doi.org/10.1016/S0010-0277(87)80006-9.

Liberman, Z., & Shaw, A. (2020). Even his friend said he's bad: Children think personal alliances bias gossip. *Cognition*, *204*, 104376. https://doi.org/10.1016/j.cognition.2020.104376.

Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, *43*, E1. https://doi.org/10.1017/S0140525X1900061X.

Liu, S., Ullman, T. D., Tenenbaum, J. B., & Spelke, E. S. (2017). Ten-month-old infants infer the value of goals from the costs of actions. *Science*, *358*(6366), 1038–1041. https://doi.org/10.1126/science.aag2132.

Lucas, C. G., Griffiths, T. L., Xu, F., Fawcett, C., Gopnik, A., Kushnir, T., et al. (2014). The child as econometrician: A rational model of preference understanding in children. *PLoS One*, *9*(3), e92160. https://doi.org/10.1371/journal.pone.0092160.

Ma, L., & Xu, F. (2011). Young children's use of statistical sampling evidence to infer the subjectivity of preferences. *Cognition*, *120*(3), 403–411. https://doi.org/10.1016/j.cognition.2011.02.003.

Martinez, J. E., Feldman, L. A., Feldman, M. J., & Cikara, M. (2021). Narratives shape cognitive representations of immigrants and immigration-policy preferences. *Psychological Science*, *32*(2), 135–152. https://doi.org/10.1177/0956797620963610.

Mascaro, O., & Sperber, D. (2009). The moral, epistemic, and mindreading components of children's vigilance towards deception. *Cognition*, *112*(3), 367–380. https://doi.org/10.1016/j.cognition.2009.05.012.

Meng, Y., & Xu, F. (2020). In S. Denison, M. Mack, Y. Xu, & B. C. Armstrong (Eds.), *How do disparities reproduce themselves? Ground truth inference from utility-maximizing agent's sampling behavior Proceedings of the 42nd annual conference of the cognitive science society* (pp. 903–908). Cognitive Science Society.

Meng, Y., & Xu, F. (2021). Naive utility calculus underlies the reproduction of disparities in social groups. In T. Fitch, C. Lamm, H. Leder, & K. Teßmar-Raible (Eds.), *Proceedings of the 43rd annual conference of the cognitive science society* (pp. 1416–1422). Cognitive Science Society.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*(5719), 255–258. https://doi.org/10.1126/science.1107621.

Pasquini, E. S., Corriveau, K. H., Koenig, M., & Harris, P. L. (2007). Preschoolers monitor the relative accuracy of informants. *Developmental Psychology*, *43*(5), 1216–1226. https://doi.org/10.1037/0012-1649.43.5.1216.

Perner, J., & Ruffman, T. (2005). Infants' insight into the mind: How deep? *Science*, *308*(5719), 214–216. https://doi.org/10.1126/science.1111656.

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology*, *33*(1), 12–21. https://doi.org/10.1037/0012-1649.33.1.12.

Rhodes, M., Gelman, S. A., & Brickman, D. (2010). Children's attention to sample composition in learning, teaching and discovery: Sample composition and learning. *Developmental Science*, *13*(3), 421–429. https://doi.org/10.1111/j.1467-7687.2009.00896.x.

Schmidt, M. F. H., Butler, L. P., Heinz, J., & Tomasello, M. (2016). Young children see a single action and infer a social norm: Promiscuous normativity in 3-year-olds. *Psychological Science*, *27*(10), 1360–1370. https://doi.org/10.1177/0956797616661182.

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, *4*(8), 299–309. https://doi.org/10.1016/S1364-6613(00)01506-0.

Schütte, F., Mani, N., & Behne, T. (2020). Retrospective inferences in selective trust. *Royal Society Open Science*, 7(2), 191451. https://doi.org/10.1098/rsos.191451.

Shafto, P., Eaves, B., Navarro, D. J., & Perfors, A. (2012). Epistemic trust: Modeling children's reasoning about others' knowledge and intent. *Developmental Science*, *15*(3), 436–447. https://doi.org/10.1111/j.1467-7687.2012.01135.x.

Shafto, P., Goodman, N. D., & Griffiths, T. L. (2014). A rational account of pedagogical reasoning: Teaching by, and learning from, examples. *Cognitive Psychology*, *71*, 55–89. https://doi.org/10.1016/j.cogpsych.2013.12.004.

Shu, T., Bhandwaldar, A., Gan, C., Smith, K. A., Liu, S., Gutfreund, D., et al. (2021). AGENT: A benchmark for core psychological reasoning. *ArXiv. 2102.12321 [Cs]* http://arxiv.org/abs/2102.12321.

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, *99*(4), 605–632. https://doi.org/10.1037/0033-295X.99.4.605.

Tomasello, M. (2016). Cultural learning redux. *Child Development*, *87*(3), 643–653. https://doi.org/10.1111/cdev.12499.

Ullman, T. D., & Tenenbaum, J. B. (2020). Bayesian models of conceptual development: Learning as building models of the world. *Annual Review of Developmental Psychology*, *2*(1), 533–558. https://doi.org/10.1146/annurev-devpsych-121318-084833.

Wellman, H. M., Kushnir, T., Xu, F., & Brink, K. A. (2016). Infants use statistical sampling to understand the psychological world. *Infancy*, *21*(5), 668–676. https://doi.org/10.1111/infa.12131.

Woodward, A. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*(1), 1–34. https://doi.org/10.1016/S0010-0277(98)00058-4.

Xu, F. (2019). Towards a rational constructivist theory of cognitive development. *Psychological Review*, *126*(6), 841–864. https://doi.org/10.1037/rev0000153.

Xu, F., & Griffiths, T. L. (2011). Probabilistic models of cognitive development: Towards a rational constructivist approach to the study of learning and development. *Cognition*, *120*(3), 299–301. https://doi.org/10.1016/j.cognition.2011.06.008.

Xu, F., & Kushnir, T. (Eds.). (2012). *Advances in child development and behavior volume 43: Rational constructivism in cognitive development*. Waltham, MA: Academic Press.

Xu, F., & Tenenbaum, J. B. (2007a). Word learning as Bayesian inference. *Psychological Review*, *114*(2), 245–272. https://doi.org/10.1037/0033-295X.114.2.245.

Xu, F., & Tenenbaum, J. B. (2007b). Sensitivity to sampling in Bayesian word learning. *Developmental Science*, *10*(3), 288–297. https://doi.org/10.1111/j.1467-7687.2007.00590.x.

This page intentionally left blank

SERIES EDITOR Jeffrey J. Lockman

ELSEVIER

ACADEMIC PRESS

An imprint of Elsevier

elsevier.com/books-and-journals

ISBN 978-0-323-99076-9

9 780323 990769