Contents lists available at ScienceDirect

Cognitive Development

journal homepage: www.elsevier.com/locate/cogdev



Simplicity and validity in infant research

Jonathan F. Kominsky ^{a,*}, Kelsey Lucca ^b, Ashley J. Thomas ^c, Michael C. Frank ^d, J. Kiley Hamlin ^e

^a Harvard Graduate School of Education, USA

^b Arizona State University, USA

^c MIT, USA

^d Stanford University, USA

^e University of British Columbia, Canada

ARTICLE INFO

Keywords: Infant Social cognition Puppets Construct validity Methods

ABSTRACT

Infancy researchers often use highly simplified, animated, or otherwise artificial stimuli to study infant's understanding of abstract concepts including "causality" or even "prosociality". The use of these simplified stimuli have led to questions about the validity of the resulting empirical findings. Do simplified stimuli effectively communicate abstract concepts to infants? Even if they do, why not use stimuli more like what infants encounter in their everyday lives? Here we make explicit the underlying logic of using simplified stimuli in studies with infants: Simplified stimuli allow for stronger experimental control and therefore more precise inferences compared to more complex, uncontrolled, naturally occurring events. We discuss the inherent tradeoff between measurement validity and ecological validity, offer three strategies for assessing the validity of stimuli to assess the development of complex social concepts in the infant mind. Ultimately, we conclude that while concerns about the validity of experiments using simplified stimuli are founded, results from such studies should not be dismissed purely on ecological grounds.

1. Introduction

One goal of developmental psychology is to understand how the mind works by examining when and how various constructs emerge during development. In recent years, many developmental researchers have increasingly asked whether young, preverbal infants possess relatively sophisticated concepts; for example, an understanding of causality (for a review see Saxe & Carey, 2006), physics (e.g., Baillargeon, 1995; Kominsky et al., 2017; Stahl & Feigenson, 2015), number (e.g., Lipton & Spelke, 2003; Xu & Spelke, 2000), the goals and intentions of agents (e.g., Csibra, 2008; Gergely, Nádasdy, Csibra, & Bíró, 1995), and even abstract social concepts such as 'helping,' 'fairness,' 'ingroup,' and 'dominance,' amongst others (e.g., Powell & Spelke, 2018; Thomsen, Frankenhuis, Ingold-Smith, & Carey, 2011; Geraci & Surian, 2011; Hamlin, Wynn, & Bloom, 2007). Results from studies like these have led some infancy researchers to conclude that preverbal infants possess rich, abstract theories of the world (e.g., Carey, 2009; Xu, 2019); others have found these rich conclusions unwarranted (e.g., Haith & Benson, 1998; Heyes, 2014; Mix, Huttenlocher, & Levine, 2002).

Notably, all of the findings purported to support the existence of early developing concepts and capabilities have come from preverbal infants who have limited attention spans, processing capacities, and fine and gross motor abilities. Indeed, studying the

* Corresponding author. *E-mail address:* jonathan.f.kominsky@gmail.com (J.F. Kominsky).

https://doi.org/10.1016/j.cogdev.2022.101213

Received 25 August 2021; Received in revised form 19 May 2022; Accepted 13 June 2022 0885-2014/@ 2022 Elsevier Inc. All rights reserved.





minds of infants necessarily requires different methods than studying the minds of adults or verbal children, even when one wishes to examine the possibility that these minds possess the very same concepts. This article focuses on one of the ways researchers attempt to address limitations inherent to infant populations: the use of highly simplified stimuli. For instance, rather than demonstrating 'causality' via one real-world object bumping into another, infants are shown animations in which one 2-dimensional circle contacts another; rather than demonstrating 'helping' via one person assisting another, infants are shown puppets or animated characters (e.g., colorful shapes with or without eyes) demonstrating and facilitating unfulfilled goals. Although such simplified stimuli are widely used throughout the infancy literature to explore infants' understanding of various constructs of interest, they are fundamentally dissimilar to infants' everyday experience with these same constructs of interest in the real world. Given this, why do researchers use simplified stimuli, and how do infants perceive them? How can we know whether and how such methods actually probe the constructs researchers intend to study? Can experiments that use these stimuli teach us anything about infants' actual mental lives?

1.1. Theoretical underpinnings of in-lab infant research

Before we venture an answer to these questions, we begin with a statement about our general perspective on psychological research, independent of whether this research uses simplified stimuli or not. Our discussion below follows the tradition of research that uses small-scale behavioral or physiological responses measured in controlled lab settings to draw inferences about human cognition (Neisser, 1976; Stevens, 1957). With regard to developmental research, we take what we would describe as a 'cognitivist' stance: we presuppose that abstract concepts appear at some point during infancy and/or early childhood, and that there is some connection between these early-developing abstractions and adult concepts. In most cases, we expect that these concepts undergo substantial changes and perhaps some degree of transformation across development (e.g., Carey, 2009; Spelke & Kinzler, 2007; Xu, 2019); the nature and extent of these changes likely differ across domains.

We acknowledge that the cognitivist approach to infancy research is not universally held. Some have argued that the concepts that we see in adults emerge from, and are only meaningfully examined during, children's interactions with the environment; that is, because adult concepts are completely discontinuous with what drives infant behavior in the lab (e.g., Allen & Bickhard, 2013; Kagan, 2008), there is intrinsically no validity to in-lab studies that isolate behavior from its naturalistic context (e.g., Packer & Moreno-Dulcy, 2020). These theories hold that the entire endeavor of infant cognitive research is largely futile in principle. Others have challenged the methods used in much of infant research, and suggested that infants' in-lab behavior can and should only be explained by purely perception-driven models that require no conceptual or representational content at all (e.g., Schöner & Thelen, 2006; Paulus, 2022).

While we acknowledge these alternative viewpoints, we do not agree with them. First, theories that deny the validity of in-lab studies in principle have trouble explaining why infants should behave in a consistent manner in the lab in any given study, much less across dozens of studies purporting to study the same conceptual content. Second, perception-driven models have sometimes proven fruitful (e.g., Bremner, Slater, Mason, Spring, & Johnson, 2013), but to date they have only been tested in a narrow subset of looking time research and fail to explain the links between infants' looking times and their adaptive, real-world behaviors (e.g., Stahl & Feigenson, 2015). The goal of the current paper is not to settle this debate, however. Our concern is whether and how using simplified stimuli rather than rich naturalistic stimuli in infant looking-time research (and other research paradigms) affects the validity of conclusions drawn from those studies. This question is wholly irrelevant to some of the alternative perspectives mentioned above, as these would reject the view that in-lab infant research can yield meaningful insight into development at all, regardless of the nature of the stimuli used. Therefore, to constrain the topics of discussion in the remainder of the paper, we begin by accepting two broadly held presuppositions in infancy and cognitive and social development research: 1) lab-based measures could, in principle, inform us about the nature of infants' minds and 2) that these insights could, in principle, translate to a better understanding of their concurrent and later ecological behavior (we return in our conclusion to whether the evidence supports #2 below).

Importantly, we do not presuppose any particular conclusion from this research concerning the relative effects of 'nature' and 'nurture' on infants' minds, and do not feel that a cognitivist approach to infancy research requires coming down on either side of the nature/nurture debate. Both nativist and empiricist viewpoints suggest that abstract concepts exist in the mind; primary disagreements concern when and how they originate (Allen & Bickhard, 2013; Haith, 2013). Indeed, the methodologies we discuss below can and have been used to test both empiricist and nativist positions; furthermore, we as an author group do not uniformly favor nativist versus empiricist conclusions regarding the nature of the infant mind.

1.2. Simplicity and validity

The goal of the current article is to discuss the validity of simplified stimuli in studies probing infant cognition, framing the issue as a trade-off between measurement and ecological validity. After outlining the validity trade-off, we focus more deeply on why simplified stimuli are sometimes needed when studying preverbal infants, and suggest three concrete strategies for assessing the validity of results generated with simplified stimuli, using illustrative examples from the literature. Finally, we assess the usefulness of these strategies within the recent and rapidly expanding literature using puppets and other simplified characters to depict agents possessing various and potentially complex mental states, whose interactions are intended to illustrate abstract and complex social constructs, including 'helping,' 'fairness,' 'hindering,' 'ingroup,' and 'dominance,' amongst others. Specifically, we ask whether researchers are truly exploring infants' emerging social cognitive capacities, ones that they might apply to the agents in their everyday lives, or instead merely examining infants' "theory of puppets" (Packer & Moreno-Dulcey, 2022). Ultimately, we conclude that when researchers keep the proposed strategies in mind, simplified stimuli can indeed be valid tools for assessing the minds of preverbal infants, including within complex social cognitive domains (see Schmuckler, 2001, for a special issue discussing similar concerns in the

context of multimodal perception research).

2. Validity in infant cognition research: a trade-off

Validity, how well a method taps into constructs of interest, can be assessed from multiple angles. *Construct validity* assesses the extent to which a measurement effectively probes the construct of interest as it exists in the participant population's everyday experience of the world (Cronbach & Meehl, 1955). Within the broader domain of construct validity, *measurement* (or *internal*) validity refers to whether observed results (whether supporting the alternative hypothesis or the null) are generated by infants' understanding of the construct that the researchers intended to study. A measure is 'measurement invalid' when there are explanations for the results that are unrelated to the construct of interest. In contrast to measurement validity, *external validity* refers to whether a measure is generalizable across contexts and populations. That is, if the measurement is done in the lab, will findings from the lab relate to how infants respond to the construct when they encounter it in the real world? And if a measure relates to particular outcomes in one population (children of one socioeconomic status, from one culture, etc.), will it also relate to those outcomes in children in another population?

Notably, in infancy research, measurement validity and external validity are often in tension with one another, creating a validity tradeoff. That is, the more stimuli used in the lab resemble how infants typically encounter a construct in the real world, the more complex the stimuli are likely to be, and the more likely they are to contain confounds that present alternative explanations for observed results and/or distractions that may prevent infants from responding systematically. To ensure that infants are responding to constructs of interest, then, infancy researchers often use highly simplified stimuli that allow the researcher to isolate the construct(s) of interest as narrowly and directly as possible. Simplified stimuli, however, are potentially far less realistic: hence the trade-off.

This more general tension between measurement and ecological validity is perhaps most prominent in the methods commonly used for exploring infants' cognitive capacities: measures of infants' relative looking time or preferential reaching (e.g., Schmuckler, 2001). In these, how long infants look at certain stimuli (e.g., Colombo & Mitchell, 2009; Rankin et al., 2009; Cruz-Khalili, Bettencourt, Kohn, Normand, & Schlinger, 2019), or whether they reach for one stimulus or another in response to previous observations (e.g., Feigenson, Carey, & Spelke, 2002; Hamlin et al., 2007) is taken to be a measure of various cognitive processes. Given there are often multiple ways to interpret these – rather generic – looking and reaching behaviors, their use brings interpretation challenges. For example, did a baby look more at an unexpected event or reach toward an object because they understood and/or evaluated the construct that researchers intended to study, or because of reactions to other features of the stimuli that happen to correlate with that construct? Because these dependent measures allow only a small number of measurements and limited comparisons between stimuli, measuring looking and reaching behavior can make it particularly difficult to know if one's stimuli are valid representations of the target construct (Aslin, 2007; Munakata, 2000; Tafreshi, Thompson, & Racine, 2014; Trehub, 2012).¹ And, even in cases where the stimuli are face-valid to some degree (e.g., language learning from recorded speech), the lab environment itself may differ from the infant's everyday environment in relevant ways (e.g., McMillan & Saffran, 2016).

The concern articulated above is a concern about ecological and measurement validity, but there are broader questions of whether looking and reaching behavior are externally valid measurements at all. That is, does looking time relate to any other measures of a given construct, either concurrently (at the same point in development) or predictively? If it does not, then one might rightly be concerned that looking time or reaching behaviors do not capture understanding of the construct at all, and no conclusions can be drawn when experimenters use these measures. A small but important literature has attempted to link individual differences in looking time measures of a variety of constructs during infancy to later measures that are related to children's adaptive behavior in more real-world contexts (e.g., Perez & Feigenson, 2021; Elliott, Feigenson, Halberda, & Libertus, 2019; Dougherty & Haith, 1997). Unfortunately this important longitudinal work is difficult to perform and generally suffers from small sample sizes and limited control variables, making strong conclusions premature. That said, based on what evidence is currently available, we adopt the assumption that in-lab looking time and reaching paradigms do in principle have some external validity, that is, they can be related to infants' capacities and behaviors outside the lab and/or later in development. However, the validity of any particular study hinges on whether its stimuli and design successfully navigate the trade-off between measurement and external validity described above, which will be the focus of our argument.

3. Why simplified stimuli? Potential confounds and extraneous information

Simplified stimuli do not perfectly reflect the real world in which infants live, and so raise external validity concerns. So, shouldn't researchers *always* prefer highly ecologically valid (e.g., real-world) stimuli? Uniformly preferring real-world over simplified stimuli can overlook serious threats to measurement validity. Real-world stimuli are complex, and complex stimuli often contain many sources of information in addition to the construct of interest. This extraneous information may undermine measurement validity in infant (or indeed any) participants, for two broad reasons.

First, extraneous information necessarily increases the risk of an unintended "confound"; that is, something in the stimuli which reliably correlates with, but is in principle unrelated to, the construct of interest. For example, if a group of researchers is trying to

¹ While recent work has expanded to include other methods like neuroimaging (e.g., fNIRS; Gervain et al., 2011) or pupillometry (e.g. Zhang, Jaffe-Dax, Wilson, & Emberson, 2019), these methods are only just starting to be widely applied, and have other challenges in their interpretation that we will not cover here.



4

Fig. 1. Some examples of minimal stimuli used in studies with infants. (a) A schematic 'launching' event with two geometric shapes, similar to the displays used by Leslie and Keeble (1987). (b) Puppets interacting with a human actor, used in Thomas, Saxe, and Spelke (2020). (c) "Helping" and "Hindering" displays with agents consisting of geometric shapes with eyes, used in Hamlin et al. (2007).

study numerical reasoning, their experimental design must account for the fact that in the real world, more versus fewer objects often correlates with more versus less of other variables such as surface area. If researchers fail to account for these correlations, then their measure may be invalid because infants' responses may stem from attending to area (a confound) rather than number (the construct of interest; see, e.g., Wynn, 1992). In these cases, multiple additional experiments are often necessary to identify the exact source of infants' responses.

The second way that extraneous information might undermine measurement validity is if it distracts infants from the target information and/or generally overwhelms infants' limited information processing capacities. In this case, infants may be unable to attend to construct-relevant aspects of the stimuli, making even those infants who understand something about a given construct fail to respond systematically (see, e.g., Tummeltshammer, Mareschal, & Kirkham, 2014). In both of these situations, the presence or absence of systematicity in infants' responses is related to the complexity of the stimuli, and unrelated to their sensitivity to the construct of interest. Notably, we are not claiming that infants cannot understand these constructs within more complex situations in their everyday lives, only that they may not be able to respond to them in a systematic way within the relatively contrived parameters of a brief laboratory testing session. Thus, studies using complex stimuli may be measurement invalid.

To better illustrate these concerns, consider the literature exploring whether and when infants perceive causality (for review, see Saxe & Carey, 2006; Carey, 2009). In an early study using simplified stimuli (Leslie & Keeble, 1987), 7-month-olds were habituated to events in which square A moved toward square B until they were adjacent, at which point A stopped and B immediately started moving (see Fig. 1a); critically, this manipulation reliably gives adults the impression of a causal collision (or 'launching') event (Michotte, 1963; Scholl & Tremoulet, 2000; for animated examples see https://www.jfkominsky.com/demos.html). At test, infants were shown B launching A, and they dishabituated to the reversal of causal roles, suggesting they had interpreted the habituation event in terms of its causal structure. The causal condition was contrasted with a non-causal condition, in which there was a 500 ms pause between the initial object making contact and the subsequent movement of the second object; this delay is known to eliminate the impression of causality in adults (Michotte, 1963; Scholl & Tremoulet, 2000). Here, Leslie and Keeble observed that infants in the non-causal condition did not dishabituate to the role reversal, suggesting that they, too, did not view the habituation event as causal.

Had Leslie and Keeble (1987) used real-world, ecologically valid events of 3-D objects colliding (for example, cars), information about motion and timing would have been accompanied by many extraneous pieces of information: multimodal cues (e.g., sight and sound), visual effects of the crash (e.g., compression), the aftermath of the crash (e.g., debris), etc. This additional information, though rich and realistic, would nevertheless introduce multiple confounds to the intended causality construct. For instance, it would be impossible to disentangle whether a difference in looking time across trial types or conditions was caused by the movement of the second object (as intended), or instead by the sound of the crash or the crushing of the car accompanying the moment of collision. As this – perhaps artificial – example illustrates, the more complex a stimulus is, the more difficult experimental control becomes.

Other studies within the causality literature also illustrate the issue of distraction. Oakes and Cohen (1990) showed 6- and 10-month-old infants causal and non-causal events involving stimuli that were relatively more ecologically valid than Leslie and Keeble (1987) abstract squares: (silent) videos of wheeled toys. With these more complex stimuli, only 10-month-olds differentiated the causal versus non-causal events, whereas 6-month-olds did not. Does that mean 6-month-olds do not understand causality, and that the Leslie and Keeble (1987) study suggesting that they do was wrong? Other work suggests not: In Cohen and Amsel (1998) replicated the Oakes and Cohen (1990) paradigm, but using simple abstract shapes (as in Leslie & Keeble, 1987) rather than complex toys; here, 6-month-olds successfully distinguished causal from non-causal events. The fact that 6-month-olds in the 1990 study failed to distinguish causal from non-causal events could lead one to conclude that causal perception for abstract shapes is unrelated to the capacity for understanding wheeled-toy and other real-world collision interactions. However, a far more parsimonious explanation in the context of the broader literature is that 6-month-olds were distracted or overwhelmed by the visual complexity of the wheeled toys themselves, and therefore their patterns of attention were based on more than just how the toys interacted with each other. By 10 months, infants are more capable of overcoming these demands on their attention in order to focus on causality. That is, it seems likely that the developmental change identified by Oakes and Cohen (1990) reflected changes in infants' processing capacities, rather than in their sensitivity to causality. Of course, removing distractions often means moving further from ecological validity, which raises the question of whether a stimulus contains the construct of interest at all. We turn to this issue in the next section.

4. Conveying construct-relevant information: three strategies

In cases where using ecologically valid stimuli would threaten measurement validity, how can we know that simplified stimuli actually represent and convey the construct of interest? Notably, getting a hypothesized result does not necessarily mean that the stimulus conveys exactly what it was intended to convey. Even with simplified stimuli, low-level perceptual features can drive infant looking behavior in systematic ways, creating potential confounds. Here we outline three strategies for assessing whether or not one's stimuli convey construct-relevant information. On its own each individual strategy has its limitations (which we will highlight below); however, we argue that convergent evidence across multiple strategies can be taken as sufficient for establishing construct validity.

First, one can attempt to validate simplified stimuli with older children or adults, who can more clearly communicate about how they view something (whether it be by overt description, task instructions whereby subjects press a button whenever they feel they have seen an instance of a construct, etc.) For example, as noted above, many decades of psychophysical research have shown that a set of precise spatiotemporal parameters leads adults to see causality in simplified interactions between geometric shapes (e.g., "launching" between two animated circles; Michotte, 1963; for recent reviews see Hubbard, 2013; Scholl & Tremoulet, 2000; for animated examples see https://www.jfkominsky.com/demos.html). Based on the assumption that infants may interpret them as causal in the same way that adults do, identical parameters have been used to create stimuli for studies of infant causal perception (see Fig. 1a;

e.g., Bélanger & Desrochers, 2001; Cohen & Amsel, 1998; Kominsky et al., 2017; Leslie & Keeble, 1987; Oakes & Cohen, 1990). Indeed, work relying on this logic has demonstrated that infants respond to manipulations of causal perception in a way that mirrors adults' responses by around 6 months of age (Cohen & Amsel, 1998). Thus, any challenge to validity would need to provide a viable alternative explanation for the developmental continuity that is observed.

All the same, comparisons with adult data are indirect, and suffer from two primary limitations. First, even if infants or children behave in a way that seems similar to how adults respond to the same stimuli, they could be doing so for entirely different reasons than adults do. For example, they could interpret the stimuli in a completely different way from adults, but their interpretation might nonetheless support the exact same looking/reaching tendencies that the interpretation based on adult responses would. Here the burden appears to fall on the creation and testing of a specific alternative explanation for a pattern of results. Second, comparisons with adult data cannot definitively address cases where infants *fail* to respond like adults or older children do. Is the lack of response due to the construct not being present early in development, or instead because the simplified stimuli only convey the construct to older children and adults by some cultural invention that infants have not yet learned about (e.g., could a 6-month-old interpret an arrow as indicating movement in a still image, as in Fig. 1a)?

A second way to establish construct validity in simplified paradigms is by gathering evidence from infant paradigms using more versus less ecologically valid stimuli and comparing their results. If results from studies using simplified stimuli resemble those from studies using real-world stimuli, it suggests that both types of stimuli are conveying similar information. A clear illustration of this strategy comes from the literature on infants' understanding of agents' goal-directed actions (see Carey, 2009, chapter 5 for a review). Within this literature, there is research examining how infants respond to ecologically valid human agents, as well as to simplified, non-human agents who nevertheless act in humanlike ways (e.g., humanoid robots, puppets, and simple shapes). Critically, aside from how agents are depicted, these studies have used nearly identical methods; therefore their results are particularly suitable for comparison. For instance, it has been shown that infants will encode the object that a person reaches for, rather than the location toward which the person reaches (e.g., Woodward, 1998), and that they expect humans to pursue their goals efficiently, by minimizing costs (e.g., Brandone & Wellman, 2009; Liu, Brooks, & Spelke, 2019). Parallel results have been found in studies using simplified non-human agents (e.g., shapes with self-generated motion): infants encode the 'goals' of non-human agents just as they do human hands (Adam, Reitenbach, & Elsner, 2017; Luo & Baillargeon, 2005), and they expect non-human agents to pursue their goals 'efficiently' just like human agents (e.g., Biro, Csibra, & Gergely, 2007; Gergely et al., 1995; Kamewari, Kato, Kanda, Ishiguro, & Hiraki, 2005). The main limitation of relying on this strategy alone is that if one fails to find convergent evidence between more ecologically-valid and more simplified stimuli, it is difficult to know whether this divergence is the result of infants being overwhelmed by real-world stimuli (as discussed above) or the result of true differences in conceptual understanding.

Finally, a third way to assess whether or not simplified stimuli convey a construct is to compare results from studies that all use simplified stimuli, but that measure more versus less ecologically valid dependent variables. For example, in the literature on infants' naive physical understanding, there is extensive evidence from looking-time studies using simplified stimuli that infants are sensitive to constraints on solidity (i.e., one solid object cannot pass through another) and gravity (i.e., an unsupported object will fall; e.g., Baillargeon, 1995; Carey, 2009; Spelke, Breinlinger, Macomber, & Jacobson, 1992); specifically, infants have been shown to look longer after viewing events in which objects violate these physical constraints. However, given inherent difficulties in interpreting results using infant looking time (Aslin, 2007; Munakata, 2000; Tafreshi et al., 2014; Trehub, 2012), such findings have not gone uncontested (Bogartz, Cashon, Cohen, Schilling, & Shinskey, 2000; Haith & Benson, 1998; Haith, 1998). Critically, convergent evidence comes from work illustrating that infants not only look longer at physically impossible events, but also adjust their subsequent actions on an object depending on the particular physical constraint it violated. In particular, Stahl and Feigenson (2015) found that if an object previously violated solidity constraints (e.g., passed through a solid wall), infants later bashed it against their high chair as though testing whether it breaks apart; if the object instead violated gravity constraints (e.g., floated in mid-air), infants instead dropped it from a height as though testing whether it falls. Under the assumption that infants are applying these actions based on past experience with how objects in general behave, this differential, targeted object exploration is far less ambiguous than differential looking time. These results more clearly support claims that the simplified physics-violating events shown to preverbal infants effectively portray the intended constructs. In turn, they suggest that findings from studies measuring infants' looking time to events containing similarly simplified stimuli are likely measurement valid. However, it is not always possible to use more ecologically or face valid responses for every construct of interest. For example, nobody has yet developed a face-valid measure of infants' causal understanding of simple collision events. In addition, given infants' physical limitations, face-valid measures may not be suitable for very young infants.

5. Challenging ecological validity in studies of complex social cognition: a "theory of puppets"?

Though such questions are certainly not new, in recent years there has been a relative explosion of research into infants' and young children's understanding of the social world (Banaji & Gelman, 2013). In addition to examining the development of relatively more basic social cognitive questions such as whether infants distinguish agents from non-agents and how individual agents are perceived and understood (for reviews see Carey, 2009; Krogh-Jesperson & Woodward, 2016), researchers have increasingly been exploring whether and when infants possess complex social concepts, those which are uniquely instantiated within the interactions between multiple agents (see e.g., Hamlin & Sitch, 2020; Powell, 2021; Ting, Dawkins, Stavans, & Baillargeon, 2019). To date, these concepts have included social goals and relationships such as "helping," "fairness," "group membership," and "dominance," amongst others, and results from studies like these have led some (but by no means all) researchers to conclude that preverbal infants possess one or more foundational systems for understanding and evaluating the social world (e.g., Geraci & Surian, 2011; Hamlin, 2013; Kinzler, Dupoux,

& Spelke, 2007; Powell & Spelke, 2013; Premack & Premack, 1997; Schmidt & Sommerville, 2011; Thomsen et al., 2011).

However, a number of important critiques of this work have emerged. Amongst others (such as mentalistic accounts of psychological reasoning versus rule-based or behavioral interpretations, e.g., Heyes, 2014; Ruffman & Perner, 2005; Scott, Richman, & Baillargeon, 2015; Sirois & Jackson, 2007), one prominent criticism is the validity issue at the center of the current paper: that because much of the evidence for high-level social capacities in preverbal infants comes from studies using highly simplified, non-human agents, the findings can tell us little about infants' understanding of the actual social worlds in which they live. That is, these studies may tell us about infants' "theory of puppets," but little else (e.g., Packer & Moreno-Dulcy, 2020; Packer & Moreno-Dulcey, 2022). In this section, we first delve into why the use of simplified agents may be particularly useful for examining infants' understanding of complex social interactions, and then we consider to what extent the three strategies for assessing measurement validity in studies using simplified stimuli outlined above can be effectively applied to this work.

5.1. Why use simplified stimuli for studying social cognition?

As discussed above, despite their lack of external validity, simplified stimuli can help to guard against threats to measurement validity, including the influence of confounding variables and/or distraction on infant's responses. Methodologies requiring the use of multiple agents are perhaps particularly susceptible to these validity threats. Indeed, a large literature demonstrates that infants exhibit various a priori preferences for people based on their gender, ethnicity, attractiveness, etc. (Bar-Haim, Ziv, Lamy, & Hodes, 2006; Kelly et al., 2005; Langlois et al., 1987; Quinn, Yahr, Kuhn, Slater, & Pascalils, 2002; Slater et al., 1998); such preferences are known to be related, at least in part, to each infants' unique experiential history (Bar-Haim et al., 2006; Quinn et al., 2002). These infant-specific baseline preferences, when unrelated to the construct of interest, could lead to spurious or null findings in studies that require infants to compare one person to another, as is often required in studies of high-level social cognition involving multiple agents.

Furthermore, despite infants' general orientation toward the social world (Johnson, Posner, & Rothbart, 1991), infants in many cultural contexts show extreme wariness of unfamiliar people starting from around 6–8 months of age (Field, 2008; Waters, Matas, & Sroufe, 1975). This wariness could prevent infants from attending or responding to the intended differences between agents, or indeed from participating in the experiment at all. Therefore, wary infants often need to be "warmed up" to any persons that they are later expected to observe and/or interact with (Cortes Barragan & Dweck, 2014; Warneken & Tomasello, 2013). However, for experimental designs requiring that infants interact with multiple human agents, this kind of warm-up procedure can be very difficult to implement effectively. For instance, multiple warm-ups may make the experimental procedure unreasonably long for parents, infants, and/or the (typically student) actors, to say nothing of the logistical challenges of scheduling all of those parties for extended blocks of time for each and every baby participating in an experiment. Further, experimenter effects – that some experimenters may simply make babies feel more or less comfortable – may introduce another source of variance that the researchers did not intend to study.²

In addition to infants' baseline likes and dislikes, studies with interacting human agents may also introduce a variety of potential experimental confounds. Consider, for example, morally relevant concepts such as helping versus harming or fair versus unfair resource distribution. Real-world prosocial and antisocial acts between humans are typically associated with multiple convergent cues, only some of which actually instantiate the construct of interest. Should the human actors smile, for example? Whereas a smile would be inappropriate in an antisocial interaction, *failing* to smile in a prosocial interaction might itself be odd. Given that infants readily distinguish different emotional expressions from early in life (Grossmann, Striano, & Friederici, 2007; Walker-Andrews, 1997) and form impressions of human agents who previously smiled versus frowned (Krol & Grossmann, 2020), it is possible that they would notice any differential facial expressions presented by human agents and use these, rather than the construct of interest, to evaluate the agents in the study. Although researchers could simply have all human experimenters maintain a neutral expression throughout their interactions (e.g., Schmidt & Sommerville, 2011), or control for the overall amount of each emotion across interaction type (e.g., Buon et al., 2014), doing so in a way that does not itself threaten validity can be difficult to achieve. For instance, it is highly unusual to see humans with still faces in the "real world," and is well-known to disturb infants during face-to-face interaction (Mesman, van LJzendoorn, & Bakermans-Kranenburg, 2009).³

5.2. Strategies for validating the use of simplified stimuli for studies of complex social cognition

Given concerns with baseline preferences and emotional confounds in studies requiring the use of multiple agents, researchers often choose to replace human actors with non-human agents such as puppets, stuffed animals, and other simplified characters such as geometric shapes with or without eyes (for various examples see Fig. 1); these simplified agents may be less susceptible to baseline preferences, and need not wear emotional expressions. But of course the use of these stimuli, and the interpretation of subsequent

² It is worth noting that this "noise" may itself be meaningful, interesting, and worthy of study. Indeed, a strongly empiricist view of development would hold that "noise" reflects meaningful variation, particularly in complex contexts like social interaction. However, presumably one must intend to study this noise and design one's methods and analyses accordingly, which most of the studies we cite in this section do not.

³ Of course, if one wanted to study the role of emotional expressions or preferences for specific individuals in infants' social evaluations, then one might in fact *want* to use these sorts of ecologically valid stimuli, but that is a different question from the one these papers were attempting to study. More broadly, there are some questions for which more ecologically valid stimuli are preferable or even necessary, but it depends heavily on what exactly the study is trying to examine. Across all of psychology, one study's confound can be another study's research question (for a particularly striking example of this in the adult literature, see Phillips et al., 2015).

results, rests on the critical assumption that infants interpret interactions between non-human agents as they interpret interactions between humans. Do they? Here we apply the three strategies outlined above for examining the validity of using simplified stimuli for studies of complex social cognition in infancy.

First, we consider adults' perceptions of simplified stimuli. Adults readily attribute complex social goals to stimuli depicting interactions between multiple simplified non-human agents. For example, classic work from Heider and Simmel (1944) showed adult participants three geometric shapes moving around on a white background. When asked to describe what they saw, adults produced highly complex mentalistic descriptions of each of the characters involved, including hypotheses about their underlying personalities and relationships. This phenomenon is so robust that you can experience it yourself simply by watching displays modeled on Heider and Simmel's own stimuli (e.g., http://perception.yale.edu/DemoFiles/Animacy-Basics/HS-Blocks-QT.mp4⁴). Other work has used similar shape-based agents to specifically identify the psychophysical parameters underlying the perception of social goals such as chasing or stalking (Gao, Newman, & Scholl, 2009); indeed, studies of the "Wolfpack Effect" have shown that adults will even spontaneously attribute animacy and purposeful movement to multiple interacting shapes that are in fact inanimate (that is, they move entirely randomly), so long as those shapes are oriented toward another randomly moving shape (Gao, McCarthy, & Scholl, 2010). These effects are apparently automatic and irresistible: adults even show them when doing so impairs performance on another task. Together, this work with adults has been taken to suggest that animacy/agency detection, including detection of social goals involving multiple agents, is very much a *perceptual* process, built into the human visual system itself (Scholl & Gao, 2013). This body of evidence provides strong support for the possibility that the simplified stimuli used to explore infants' understanding of high-level social concepts reliably portrays those concepts to adult participants, consistent with our first strategy.

In addition to adults' interpretation of shape-based agents per se, the broader literature on anthropomorphism (Epley, Waytz, & Cacioppo, 2007; Waytz, Epley, & Cacioppo, 2010) provides strong convergent evidence for the possibility that infants apply social reasoning to non-human agents. This literature demonstrates that adults (and children; Piaget, 1929; Severson & Lemm, 2016) readily engage in rich social reasoning about non-human entities as diverse as religious deities, pets, cartoon characters in books and movies, robots, and even to our more autonomous household appliances (e.g., Sung, Guo, Grinter, & Christensen, 2007), imbuing them with a host of intentional, motivational, and emotional states. The tendency to anthropomorphize is often considered an automatic feature of typical human psychology (Guthrie, 1997). Indeed, rather than something that adults are particularly likely to do as the result of a lifetime of practice applying social reasoning to human agents, developmental evidence suggests that children are relatively *more* likely than adults to anthropomorphize (e.g., Bering & Bjorklund, 2004), perhaps because children have relatively fewer cognitive resources available to override the automatic process that causes anthropomorphic thinking in the first place (see e.g., Epley, Morewedge, & Keysar, 2004). Thus, the broader literature on anthropomorphism suggests that typically developing humans, but perhaps particularly young children, readily apply social reasoning to non-human agents just like those utilized in studies of complex social cognition in infancy. This convergent evidence supports the possibility that infants apply social reasoning to one-human agents just like those utilized in studies of complex social cognition in infancy. This convergent evidence supports the possibility that infants apply social reasoning to one-human agents just like those utilized in studies of complex social cognition in infancy. This convergent evidence supports the possibility that infants apply social reaso

The second strategy for assessing the validity of infant studies using simplified agents is to compare results from studies using simplified agents to those from studies using human agents. For instance, the first studies of infants' evaluation of helping and harming utilized shape-based agents with eyes, and demonstrated that infants selectively approached and/or looked longer toward helping shapes than toward hindering shapes, (e.g., Hamlin, 2014; Hamlin et al., 2007; Hamlin, Wynn, & Bloom, 2010; Lee, Yun, Kim, & Song, 2015; c.f. Scarf, Imuta, Colombo, & Hayne, 2012; Schlingloff, Csibra, & Tatone, 2020; see also Geraci & Simion, 2021; Kanakogi et al., 2017; Margoni & Surian, 2018). Subsequent studies found convergent evidence utilizing other types of helpful and harmful agents; for instance, infants also selectively approached animal puppets who helped versus hindered (Hamlin & Wynn, 2011; Hamlin, 2013; Steckler, Woo, & Hamlin, 2017; c.f. Salvadori et al., 2015), cartoon humans who gave versus took (Scola, Holvoet, Arciszewski, & Picard, 2015), and real humans who aggressed against an object and comforted another human versus aggressed against another human and comforted an object (Buon et al., 2014). Similarly within the fairness domain, infants have been shown to respond similarly to studies using a variety of types of agents, both looking longer following unfair versus fair distributions (e.g., Buyukozer-Dawkins, Sloane, & Baillargeon, 2019; Meristo, Strid, & Surian, 2016; Schmidt & Sommerville, 2011) and preferring fair to unfair agents (e.g., Burns & Sommerville, 2014; Geraci & Surian, 2011; Lucca, Pospisil, & Sommerville, 2018; Surian & Franchin, 2017). Critically, infants show these tendencies whether the fair and unfair agents were humans, animal puppets, or shapes. Finally, a recent meta-analysis of over 250 studies suggests that performance on theory-of-mind tasks is exactly the same regardless of whether the agents are puppets or people (Yu & Wellman, in press). Thus, at least for those social concepts for which both human and non-human agents have been used. the evidence to date suggests that infants respond similarly to human and non-human agents. These convergent results support the possibility that studies using simplified agents are no less valid than those using human actors.

The third and final strategy for assessing the validity of simplified stimuli is to compare results from studies examining more- and less- ecologically-valid dependent variables. Indeed, in addition to looking time and preferential approach, infants appear to treat simplified agents in at least some face-valid ways, engaging in a variety of interactive behaviors with them. For instance, infants have been shown to imitate non-human puppets as early as 6 months of age (Barr, Dowden, & Hayne, 1996), and by 15–18 months infants enact the unfulfilled (and unseen) goals of both human and non-human (e.g., stuffed animal) agents (e.g., with humans: Meltzoff, 1995; with stuffed animals: Johnson, Booth, & O'Hearn, 2001). Critically, infants in these same studies do not enact unseen endstates of physically similar actions produced by inanimate entities, suggesting that imitation is selective to (human and non-human) agents. Other studies with non-human agents suggest that infants will readily follow non-human agents "gaze," perhaps because they consider

⁴ This demo animation created by Brian Scholl and referenced with gratitude.

gaze as signaling communicative intent: 12-month-olds look in the direction a faceless blob looked, so long as the blob had previously acted contingently with the infant or with another agent (e.g., Johnson, Slaughter, & Carey, 1998), and 18-month-olds reliably followed the gaze of a robot, if and only if the robot had previously engaged in a contingent social interaction with a human (Meltzoff, Brooks, Shon, & Rao, 2010). Notably, in the latter study infants did not gaze-follow a non-contingent robot, even though the robot had a face with eyes. These results suggest that (although faces can act as agency cues in some contexts) it is acting as though one is able to perceive the world, as opposed to the possession of a face, that influences whether or not infants gaze-follow.

Finally, toddlers have also been shown to engage in longer and/or more complex types of interactions with non-human agents, suggesting that they treat them as meaningful interaction partners under some circumstances. For example, in one study 17-montholds "helped" a shape-based agent in its demonstrated unfulfilled goal to climb over a barrier (Kenward & Gredebäck, 2013). Critically, and similarly to no-need control conditions in studies exploring toddlers helping human agents (e.g., Warneken & Tomasello, 2006), toddlers only lifted the agent over the barrier when the constraints of the situation suggested that the agent was actually in need. Other studies have shown that toddlers readily engage in a variety of game-like interactions with non-human agents, including giving objects and taking them away (e.g., Van de Vondervoort, Aknin, Kushnir, Slevinsky, & Hamlin, 2018; see also Kenward & Dahl, 2011). Indeed, toddlers in both North America and the remote island nation of Vanuatu appear to reap emotional benefits (e.g., look happy) after sharing with stuffed animals (e.g., Aknin, Hamlin, & Dunn, 2012; Aknin, Broesch, Hamlin, & Van de Vondervoort, 2015), suggestive that their prosocial interactions with non-human agents can be emotionally meaningful. Together, results from studies like these strongly indicate that infants and toddlers are capable of treating non-human agents as true social agents, and interact with them as such.

Taken together, these areas of related research suggest that infants indeed interpret simplified social stimuli as social. But importantly, our arguments should not be taken to imply that researchers yet know exactly how infants represent the social agents they encounter within these studies. For example, Powell and Spelke (2018) asked whether infants' preferences for helpers could be explained by a preference for imitators instead. Although their experiment did not rule out the helper hypothesis completely, this is a viable and important alternative. It is up to future research to distinguish between these possibilities. What we do argue is that studies utilizing simplified non-human agents to explore complex social cognition in infants and very young children are valid, in that it is unlikely that researchers in this field are tapping a theory of puppets or schematic stimuli that is distinct from a theory of human agents. Instead, their results can be appropriately interpreted as providing (heretofore incomplete) evidence for a theory of social agents more generally, of which humans might (in some cases) be a further specialized subset.

6. Generalizability of simplified stimuli across cultures

Another facet of ecological validity is whether a finding generalizes not just across stimuli, but across populations. Most experimental infant psychology research to date has been conducted on a non-representative sample in terms of race, ethnicity, language spoken, nationality, socio-economic status, and culture (Henrich, Heine, & Norenzayan, 2010; Nielsen, Haun, Kärtner, & Legare, 2017). By six or eight or ten months of age, there is already evidence for many cross-cultural differences (e.g., Camras et al., 1998; De Boysson-Bardies, Sagart, & Durand, 1984; Geangu et al., 2016) as well as similarities (e.g., ManyBabies Consortium, 2020). It is not always clear what the bases of these differences are, but one question that arises is the extent to which infants in different cultures interpret simplified stimuli in the same way. That is, when an experiment that uses the kind of simplified stimuli we have discussed throughout this paper is run in two (or more) different cultures, one potential threat to the ecological validity of those stimuli concerns whether they will be understood in the same way by each population.

There are some types of stimuli that we would expect to work perfectly well in any human culture. For example, the cues that identify animate agents do not depend heavily on cultural context (Carey, 2009; Rakison & Poulin-Dubois, 2001): anywhere on the planet, entities that are capable of self-propelled movement, have faces, and interact contingently are going to be animate agents. At a bare minimum, the presence of other human beings is, by definition, a cultural universal.

Furthermore, the use of depictions or illustrations of agents may also be less prone to cross-cultural effects than one might intuitively think: Even newborns respond more to schematic depictions of faces than to alternatives, and their responses appear at least as large to the schematic depictions of faces as to photos of actual faces (Farroni et al., 2005). DeLoache, Pierroutsakos, Uttal, Rosengren, and Gottlieb (1998) also reported recognition-like manual behaviors relating to pictures for 9-month-olds in both the US and Ivory Coast, suggesting at least some degree of cross-cultural consistency. Drawings appear to be perceived as easily as pictures, as well; one data point comes from Hochberg and Brooks (1962), who deprived their child of access to depictions until 18 months and nevertheless found he was fully able to name schematic line drawings when they were presented. Further, as cited above, one line of work has shown identical effects in children interacting with stuffed toys in North America and in Vanuatu (Aknin et al., 2012, 2015).

All that said, infants in different cultures, or even different households within cultures, undoubtedly have varying degrees of exposure to representational toys such as dolls and puppets, and different levels of experience with animations. Empirically, it is still largely an open question whether or how experience with representational toys and animations affects how young children and infants interpret schematic stimuli. The one key exception is a recent meta-analysis in this very issue which finds consistency in false belief paradigms using both puppets and people as stimuli across 36 different countries (Yu & Wellman, in press). However, even if a study using simplified stimuli fails to replicate across cultures, there is much to be learned by asking why. Investigating whether the stimuli convey the same information across cultures could both provide insight into how simplified stimuli are interpreted in general, and create the opportunity to ask whether a suitable stimulus that *does* convey the same information in a more culturally relevant way generates the same results. Alternatively, if one can find positive evidence that the stimuli are being interpreted the same way by two different populations even though they generate different patterns of behavior, it could indicate a meaningful cross-cultural difference

J.F. Kominsky et al.

in the construct itself. Thus, failures to replicate a specific finding across cultures do not necessarily undercut the conclusions of the original studies, but rather present valuable opportunities to better understand both the constructs that the study originally intended to examine and the processes by which the stimuli are understood.

7. Conclusions

One of the running themes of infant research is that infant minds are much more advanced than their everyday behavior implies. While some phenomena can be readily studied using ecologically valid stimuli and face-valid methods, for many cognitive capacities this is either inappropriate for reasons of measurement validity, or simply impossible for reasons of practicality. A 6-month-old infant cannot readily tell us they think one person is nice and another is mean, but with these carefully-designed studies - which often use simplified stimuli - we can show that they differentiate prosocial and antisocial individuals nonetheless. However, the validity of these studies should always be subject to critical examination. Future research could change our entire understanding of existing evidence by presenting alternate explanations for these results that make specific, testable predictions, and demonstrating that those predictions are accurate. Given the scope of the challenges that have already been addressed, we would hardly be disappointed, because such research would likely provide novel insights into the mind, uncovering previously unknown capacities or mechanisms.

Acknowledgements

We would like to thank many members of the Cognitive Development Society listserv, initiated by Martin Packer, for the discussion that inspired this paper. AJT was supported by a fellowship from the Andrew W. Mellon Foundation. The authors would also like to thank members of the Harvard Lab for Developmental Studies, the ASU Emerging Minds Lab, the UBC Centre for Infant Cognition, and the Stanford Language and Cognition Lab for feedback and discussion of this paper.

References

- Adam, M., Reitenbach, I., & Elsner, B. (2017). Agency cues and 11-month-olds' and adults' anticipation of action goals. Cognitive Development, 43, 37–48. https://doi.org/10.1016/j.cogdev.2017.02.008
- Aknin, L. B., Broesch, T., Hamlin, J. K., & Van de Vondervoort, J. W. (2015). Prosocial behavior leads to happiness in a small-scale rural society. Journal of Experimental Psychology: General, 144(4), 788–795. https://doi.org/10.1037/xge0000082
- Aknin, L. B., Hamlin, J. K., & Dunn, E. W. (2012). Giving leads to happiness in young children. PLoS One, 7(6), Article e39211. https://doi.org/10.1371/journal. pone.0039211
- Allen, J. W. P., & Bickhard, M. H. (2013). Stepping off the pendulum: Why only an action-based approach can transcend the nativist–empiricist debate. *Cognitive Development*, 28(2), 96–133. https://doi.org/10.1016/j.cogdev.2013.01.002

Aslin, R. N. (2007). What's in a look. Developmental Science, 10(1), 48-53. https://doi.org/10.1111/j.1467-7687.2007.00563.x

Baillargeon, R. (1995). A model of physical reasoning in infancy. In C. Roveer-Collier & L. P. Lipsitt (Eds.), Advances in Infancy Research (Vol. 9, pp. 305–371). Retrieved from Z3988.

Banaji, M. R., & Gelman, S. A. (Eds.). (2013). Navigating the social world: What infants, children, and other species can teach us. Oxford University Press.

Bar-Haim, Y., Ziv, T., Lamy, D., & Hodes, R. M. (2006). Nature and nurture in own-race face processing. Psychological Science, 17(2), 159–163.

- Barr, R., Dowden, A., & Hayne, H. (1996). Developmental changes in deferred imitation by 6-to 24-month-old infants. Infant Behavior and Development, 19(2), 159–170.
- Bélanger, N. D., & Desrochers, S. (2001). Can 6-month-old infants process causality in different types of causal events? *British Journal of Developmental Psychology*, *19* (1), 11–21.
- Bering, J. M., & Bjorklund, D. F. (2004). The natural emergence of reasoning about the afterlife as a developmental regularity. *Developmental Psychology*, 40(2), 217–233. https://doi.org/10.1037/0012-1649.40.2.217
- Biro, S., Csibra, G., & Gergely, G. (2007). The role of behavioral cues in understanding goal-directed actions in infancy. *Progress in Brain Research*, 164, 303–322. https://doi.org/10.1016/S0079-6123(07)64017-5

Bogartz, R. S., Cashon, C. H., Cohen, L. B., Schilling, T. H., & Shinskey, J. L. (2000). Reply to Baillargeon, Aslin, and Munakata. Infancy, 1(4), 479–490.

- Brandone, A. C., & Wellman, H. M. (2009). You can't always get what you want: Infants understand failed goal-directed actions. *Psychological Science*, 20(1), 85–91.
 Bremner, J. G., Slater, A. M., Mason, U. C., Spring, J., & Johnson, S. P. (2013). Trajectory perception and object continuity: effects of shape and color change on 4-month-olds' perception of object identity. *Developmental Psychology*, 49(6), 1021–1026. https://doi.org/10.1037/a0029398
- Buon, M., Jacob, P., Margules, S., Brunet, I., Dutat, M., Cabrol, D., & Dupoux, E. (2014). Friend or foe? Early social evaluation of human interactions. *PLoS One*, *9*(2), Article e88612. https://doi.org/10.1371/journal.pone.0088612
- Burns, M. P., & Sommerville, J. A. (2014). "I pick you": The impact of fairness and race on infants' selection of social partners. Frontiers in Psychology, 5, 93. https://doi.org/10.3389/fpsyg.2014.00093
- Buyukozer-Dawkins, M., Sloane, S., & Baillargeon, R. (2019). Do infants in the first year of life expect equal resource allocations. *Frontiers in Psychology*, 10, 116. https://doi.org/10.3389/fpsyg.2019.00116
- Camras, L. A., Oster, H., Campos, J., Campos, R., Ujiie, T., Miyake, K., & Meng, Z. (1998). Production of emotional facial expressions in European American, Japanese, and Chinese infants. *Developmental Psychology*, 34(4), 616–628. https://doi.org/10.1037/0012-1649.34.4.616

Carey, S. (2009). The origin of concepts. Oxford; New York: Oxford University Press.

- Cohen, L. B., & Amsel, G. (1998). Precursors to infants' perception of the causality of a simple event. Infant Behavior and Development, 21(4), 713–731. https://doi.org/10.1016/S0163-6383(98)90040-6
- Colombo, J., & Mitchell, D. W. (2009). Infant visual habituation. Neurobiology of Learning and Memory, 92(2), 225–234. https://doi.org/10.1016/j.nlm.2008.06.002 Cortes Barragan, R., & Dweck, C. S. (2014). Rethinking natural altruism: Simple reciprocal interactions trigger children's benevolence. Proceedings of the National Academy of Sciences, 111(48), 17071–17074. https://doi.org/10.1073/pnas.1419408111

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. Psychological Bulletin, 52(4), 281.

- Cruz-Khalili, A., Bettencourt, K., Kohn, C. S., Normand, M. P., & Schlinger, H. D. (2019). Use of Repeated Within-Subject Measures to Assess Infants' Preference for Similar Others. Frontiers in Psychology, 10, 2239. https://doi.org/10.3389/fpsyg.2019.02239
- Csibra, G. (2008). Goal attribution to inanimate agents by 6.5-month-old infants. Cognition, 107(2), 705–717. https://doi.org/10.1016/j.cognition.2007.08.001
- De Boysson-Bardies, B., Sagart, L., & Durand, C. (1984). Discernible differences in the babbling of infants according to target language. *Journal of Child Language*, 11 (1), 1–15. https://doi.org/10.1017/s0305000900005559
- DeLoache, J. S., Pierroutsakos, S. L., Uttal, D. H., Rosengren, K. S., & Gottlieb, A. (1998). Grasping the nature of pictures. Psychological Science, 9(3), 205–210. https://doi.org/10.1111/1467-9280.00039

- Dougherty, T. M., & Haith, M. M. (1997). Infant expectations and reaction time as predictors of childhood speed of processing and IQ. *Developmental Psychology*, 33(1), 146–155. https://doi.org/10.1037/0012-1649.33.1.146
- Elliott, L., Feigenson, L., Halberda, J., & Libertus, M. E. (2019). Bidirectional, longitudinal associations between math ability and approximate number system precision in childhood. Journal of Cognition and Development, 20(1), 56–74. https://doi.org/10.1080/15248372.2018.1551218
- Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. Journal of Experimental Social Psychology, 40(6), 760–768. https://doi.org/10.1016/j.jesp.2004.02.002

Epley, N., Waytz, A., & Cacioppo, J. (2007). On seeing human: A three-factor theory of anthropomorphism. Psychological Review, 114(4), 864-886.

Farroni, T., Johnson, M. H., Menon, E., Zulian, L., Faraguna, D., & Csibra, G. (2005). Newborns' preference for face-relevant stimuli: Effects of contrast polarity. Proceedings of the National Academy of Sciences, 102(47), 17245–17250.

Feigenson, L., Carey, S., & Spelke, E. (2002). Infants' discrimination of number vs. continuous extent. Cognitive Psychology, 44(1), 33–66. https://doi.org/10.1006/ cogp.2001.0760

Field, T. (2008). Problems in infancy. In M. Hersen & A. M. Gross (Eds.), Handbook of Clinical Psychology (pp. 966-1011).

- Gao, T., McCarthy, G., & Scholl, B. J. (2010). The wolfpack effect. Perception of animacy irresistibly influences interactive behavior. Psychological Science, 21(12), 1845–1853. https://doi.org/10.1177/0956797610388814
- 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
- Geangu, E., Ichikawa, H., Lao, J., Kanazawa, S., Yamaguchi, M. K., Caldara, R., & Turati, C. (2016). Culture shapes 7-month-olds' perceptual strategies in

discriminating facial expressions of emotion. Current Biology, 26(14), R663-4. https://doi.org/10.1016/j.cub.2016.05.072

- Geraci, A., & Simion, F. (2021). Evaluation of prosocial actions performed by dynamic shapes at 17 months of age. European Journal of Developmental Psychology, 1–25. https://doi.org/10.1080/17405629.2021.1957823
- Geraci, A., & Surian, L. (2011). The developmental roots of fairness: Infants' reactions to equal and unequal distributions of resources. *Developmental Science*, 14(5), 1012–1020. https://doi.org/10.1111/j.1467-7687.2011.01048.x
- Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. Cognition, 56(2), 165-193.
- Gervain, J., Mehler, J., Werker, J. F., Nelson, C. A., Csibra, G., Lloyd-Fox, S., & Aslin, R. N. (2011). Near-infrared spectroscopy: A report from the McDonnell infant methodology consortium. Developmental Cognitive Neuroscience, 1, 22–46. https://doi.org/10.1016/j.dcn.2010.07.004
- Grossmann, T., Striano, T., & Friederici, A. D. (2007). Developmental changes in infants' processing of happy and angry facial expressions: a neurobehavioral study. Brain and Cognition, 64(1), 30–41. https://doi.org/10.1016/j.bandc.2006.10.002
- Guthrie, S. E. (1997). Anthropomorphism: A definition and a theory. In R. W. Mitchell, N. S. Thompson, & H. L. Miles (Eds.), Anthropomorphism, anecdotes, and animals (pp. 50–58). State University of New York Press.

Haith, M. M. (1998). Who put the cog in infant cognition? Is rich interpretation too costly. Infant Behavior and Development, 21(2), 167-179.

- Haith, M. M. (2013). Emergent constructivism has its place—Among other possibilities. Commentary on "Stepping off the pendulum: Why only an action-based approach can transcend the nativist–empiricist debate" by J. Allen and M. Bickhard. Cognitive Development, 28(2), 144–147. https://doi.org/10.1016/j. cogdev.2013.01.005
- Haith, M. M., & Benson, J. B. (1998). Infant cognition. In W. Damon (Ed.), Handbook of Child Psychology, Vol. 2: Cognition, perception, and language (pp. 199–254). John Wiley & Sons Inc.
- Hamlin, J. K. (2013). Moral judgment and action in preverbal infants and toddlers. Current Directions in Psychological Science, 22(3), 186–193. https://doi.org/ 10.1177/0963721412470687
- Hamlin, J. K. (2014). The case for social evaluation in preverbal infants: gazing toward one's goal drives infants' preferences for Helpers over Hinderers in the hill paradigm. *Frontiers in Psychology*, *5*, 1563. https://doi.org/10.3389/fpsyg.2014.01563
- Hamlin, J.K., & Sitch, M. (2020). Understanding and evaluating the moral world in infancy. In C. Tamis-Lemonda & J. Lockman (Eds.), Handbook of Infant Development.
- 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. (2010). Three-month-olds show a negativity bias in their social evaluations. *Developmental Science*, *13*(6), 923–929. https://doi.org/10.1111/j.1467-7687.2010.00951.x

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. The American Journal of Psychology, 57(2), 243-259.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world. Behavioral and Brain Sciences, 33(2-3), 61-83. https://doi.org/10.1017/ S0140525X0999152X

Heyes, C. (2014). False belief in infancy: a fresh look. Developmental Science, 17(5), 647-659. https://doi.org/10.1111/desc.12148

- Hochberg, J., & Brooks, V. (1962). Pictorial recognition as an unlearned ability: A study of one child's performance. *American Journal of Psychology*, 75(4), 624–628. Hubbard, T. L. (2013). Phenomenal causality I: Varieties and variables. *Axiomathes*, 23(1), 1–42. https://doi.org/10.1007/s10516-012-9198-8
- Johnson, M. H., Posner, M. I., & Rothbart, M. K. (1991). Components of visual orienting in early infancy: Contingency learning, anticipatory looking, and disengaging. Journal of Cognitive Neuroscience, 3(4), 335–344. https://doi.org/10.1162/jocn.1991.3.4.335
- Johnson, S., Slaughter, V., & Carey, S. (1998). Whose gaze will infants follow? The elicitation of gaze-following in 12-month-olds. Developmental Science, 1(2), 233-238.

Johnson, S. C., Booth, A., & O'Hearn, K. (2001). Inferring the goals of a nonhuman agent. Cognitive Development, 16(1), 637-656.

Kagan, J. (2008). In defense of qualitative changes in development. Child Development, 79(6), 1606–1624. https://doi.org/10.2307/27563578

Kamewari, K., Kato, M., Kanda, T., Ishiguro, H., & Hiraki, K. (2005). Six-and-a-half-month-old children positively attribute goals to human action and to humanoidrobot motion. Cognitive Development, 20(2), 303–320. https://doi.org/10.1016/j.cogdev.2005.04.004

Kanakogi, Y., Inoue, Y., Matsuda, G., Butler, D., Hiraki, K., & Myowa-Yamakoshi, M. (2017). Preverbal infants affirm third-party interventions that protect victims from aggressors. Nature Human Behaviour, 1(2). https://doi.org/10.1038/s41562-016-0037

- Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Gibson, A., Smith, M., ... Pascalis, O. (2005). Three-month-olds, but not newborns, prefer own-race faces. Developmental Science, 8(6), F31–F36.
- Kenward, B., & Dahl, M. (2011). Preschoolers distribute scarce resources according to the moral valence of recipients' previous actions. *Developmental Psychology*, 47 (4), 1054–1064. https://doi.org/10.1037/a0023869

Kenward, B., & Gredebäck, G. (2013). Infants help a non-human agent. PLoS One, 8(9), Article e75130. https://doi.org/10.1371/journal.pone.0075130

- Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences*, 104(30), 12577–12580. (https://www.pnas.org/content/104/30/12577.short).
- Kominsky, J. F., Strickland, B., Wertz, A. E., Elsner, C., Wynn, K., & Keil, F. C. (2017). Categories and constraints in causal perception. Psychological Science, 28(11), 1649–1662. https://doi.org/10.1177/0956797617719930
- Krogh-Jesperson, S., & Woodward, A. L. (2016). Infant origins of social cognition. In L. Balter, & C. Tamis-LeMonda (Eds.), Child Psychology: A Handbook of Contemporary Issues (Third ed., pp. 105–121). New York, NY: Routledge.
- Krol, K. M., & Grossmann, T. (2020). Impression formation in the human infant brain. Cerebral Cortex Communications, 1, tgaa070. https://doi.org/10.1093/texcom/ tgaa070
- Langlois, J. H., Roggman, L. A., Casey, R. J., Ritter, J. M., Rieser-Danner, L. A., & Jenkins, V. Y. (1987). Infant preferences for attractive faces: Rudiments of a stereotype. Developmental Psychology, 23(3), 363–369. https://doi.org/10.1037/0012-1649.23.3.363
- Lee, Y.-E., Yun, J.-E. E., Kim, E. Y., & Song, H.-J. (2015). The development of infants' sensitivity to behavioral intentions when inferring others' social preferences. *PLoS One*, 10(9), Article e0135588. https://doi.org/10.1371/journal.pone.0135588

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 Lipton, J. S., & Spelke, E. S. (2003). Origins of number sense: Large-number discrimination in human infants. *Psychological Science*, 14(5), 396–401.

Liu, S., Brooks, N. B., & Spelke, E. S. (2019). Origins of the concepts cause, cost, and goal in prereaching infants. Proceedings of the National Academy of Sciences. , Article 201904410. https://doi.org/10.1073/pnas.1904410116

Lucca, K., Pospisil, J., & Sommerville, J. A. (2018). Fairness informs social decision making in infancy. PLoS One, 13(2), Article e0192848. https://doi.org/10.1371/ journal.pone.0192848

Luo, Y., & Baillargeon, R. (2005). Can a self-propelled box have a goal? Psychological reasoning in 5-month-old infants. Psychological Science, 16(8), 601–608. https://doi.org/10.1111/j.1467-9280.2005.01582.x

ManyBabies Consortium. (2020). Quantifying sources of variability in infancy research using the infant-directed-speech preference. Advances in Methods and Practices in Psychological Science, 3(1), 24–52. https://doi.org/10.1080/0144341870070303

Margoni, F., & Surian, L. (2018). Infants' evaluation of prosocial and antisocial agents: A meta-analysis. Developmental Psychology, 54(8), 1445–1455. https://doi.org/ 10.1037/dev0000538

McMillan, B. T., & Saffran, J. R. (2016). Learning in complex environments: The effects of background speech on early word learning. Child Development, 87(6), 1841–1855. https://doi.org/10.1111/cdev.12559

Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. Developmental Psychology, 31(5), 838–850. https://doi.org/10.1037/0012-1649.31.5.838

Meltzoff, A. N., Brooks, R., Shon, A. P., & Rao, R. P. (2010). "Social" robots are psychological agents for infants: A test of gaze following. Neural Networks, 23(8–9), 966–972. https://doi.org/10.1016/j.neunet.2010.09.005

Meristo, M., Strid, K., & Surian, L. (2016). Preverbal infants' ability to encode the outcome of distributive actions. Infancy, 21(3), 353–372. https://doi.org/10.1111/ infa.12124

Mesman, J., van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2009). The many faces of the Still-Face Paradigm: A review and meta-analysis. Developmental Review, 29(2), 120–162. https://doi.org/10.1016/j.dr.2009.02.001

Michotte, A. (1963). The perception of causality. New York, NY: Basic Books.

Mix, K. S., Huttenlocher, J., & Levine, S. C. (2002). Multiple cues for quantification in infancy: Is number one of them. Psychological Bulletin, 128(2), 278.

Munakata, Y. (2000). Challenges to the violation-of-expectation paradigm: Throwing the conceptual baby out with the perceptual processing bathwater. *Infancy*, 1(4), 471–477.

Neisser, U. (1976). Cognition and Reality: Principles and Implications of Cognitive Psychology. W. H. Freeman.

Nielsen, M., Haun, D., Kärtner, J., & Legare, C. H. (2017). The persistent sampling bias in developmental psychology: A call to action. Journal of Experimental Child Psychology, 162, 31–38. https://doi.org/10.1016/j.jecp.2017.04.017

Dakes, L. M., & Cohen, L. B. (1990). Infant perception of a causal event. Cognitive Development, 5(2), 193-207.

Packer, M., & Moreno-Dulcy, F.A. (2020). "This puppet will play a game with you": Is it time to take child psychology out of the laboratory? Unpublished manuscript. Packer, M. J., & Moreno-Dulcey, F. A. (2022). Theory of puppets?: A critique of the use of puppets as stimulus materials in psychological research with young children. Committive Development, 61, Article 101146.

Paulus, M. (2022). Should infant psychology rely on the violation-of-expectation method? Not anymore. Infant and Child Development, 31(1). https://doi.org/10.1002/ icd.2306

Perez, J., & Feigenson, L. (2021). Stable individual differences in infants' responses to violations of intuitive physics. Proceedings of the National Academy of Sciences, 118(27), Article e2103805118.

Phillips, J., Ong, D. C., Surtees, A. D., Xin, Y., Williams, S., Saxe, R., & Frank, M. C. (2015). A second look at automatic theory of mind: Reconsidering kovács, téglás, and endress (2010). Psychological Science, 26(9), 1353–1367. https://doi.org/10.1177/0956797614558717

Piaget, J. (1929). The Child's conception of the world. London: Routledge and Kegan Paul, Ltd.

Powell, L. (2021). Adopted utility calculus: Origins of a concept of social affiliation. (https://psyarxiv.com/kuwgf/download?format=pdf).

Powell, L. J., & Spelke, E. S. (2013). Preverbal infants expect members of social groups to act alike. Proceedings of the National Academy of Sciences, 110(41), E3965–72. https://doi.org/10.1073/pnas.1304326110

Powell, L. J., & Spelke, E. S. (2018). Third-party preferences for imitators in preverbal infants. Open Mind, 2, 61–71. https://doi.org/10.1162/opmi_a_00018

Premack, D., & Premack, A. J. (1997). Infants attribute value± to the goal-directed actions of self-propelled objects. Journal of Cognitive Neuroscience, 9(6), 848–856. https://doi.org/10.1162/jocn.1997.9.6.848

Quinn, P. C., Yahr, J., Kuhn, A., Slater, A. M., & Pascalils, O. (2002). Representation of the gender of human faces by infants: A preference for female. *Perception*, 31(9), 1109–1121. https://doi.org/10.1068/p3331

Rakison, D. H., & Poulin-Dubois, D. (2001). Developmental origin of the animate-inanimate distinction. Psychological Bulletin, 127(2), 209.

Rankin, C. H., Abrams, T., Barry, R. J., Bhatnagar, S., Clayton, D. F., Colombo, J., & Thompson, R. F. (2009). Habituation revisited: An updated and revised description of the behavioral characteristics of habituation. *Neurobiology of Learning and Memory*, 92(2), 135–138. https://doi.org/10.1016/j.nlm.2008.09.012

Ruffman, T., & Perner, J. (2005). Do infants really understand false belief? Trends in Cognitive Sciences, 9(10), 462–463. https://doi.org/10.1016/j.tics.2005.08.001
Salvadori, E., Blazsekova, T., Volein, A., Karap, Z., Tatone, D., Mascaro, O., & Csibra, G. (2015). Probing the strength of infants' preference for helpers over hinderers:
Two replication attempts of hamlin and wynn (2011). PLoS One, 10(11), Article e0140570. https://doi.org/10.1371/journal.pone.0140570

Saxe, R., & Carey, S. (2006). The perception of causality in infancy. Acta Psychologica, 123(1-2), 144–165. https://doi.org/10.1016/j.actpsy.2006.05.005

Scarf, D., Imuta, K., Colombo, M., & Hayne, H. (2012). Social evaluation or simple association? Simple associations may explain moral reasoning in infants. PLoS One, 7(8), Article e42698. https://doi.org/10.1371/journal.pone.0042698

Schlingloff, L., Csibra, G., & Tatone, D. (2020). Do 15-month-old infants prefer helpers? A replication of Hamlin et al. (2007). Royal Society Open Science, 7(4), Article 191795. https://doi.org/10.1098/rsos.191795

Schmidt, M. F. H., & Sommerville, J. A. (2011). Fairness expectations and altruistic sharing in 15-month-old human infants. PLoS One, 6(10), Article e23223. https:// doi.org/10.1371/journal.pone.0023223

Schmuckler, M. A. (2001). What is ecological validity? A dimensional analysis. Infancy, 2(4), 419-436. https://doi.org/10.1207/S15327078IN0204_02

Scholl, B. J., & Gao, T. (2013). Perceiving animacy and intentionality: Visual processing or higher-level judgment. In M. D. Rutherford, & V. A. Kuhlmeier (Eds.), Social perception: Detection and interpretation of animacy, agency, and intention (pp. 197–230). Cambridge, MA: MIT Press.

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. Trends in Cognitive Science, 4(8), 299–309. https://doi.org/10.1016/S1364-6613(00)01506-0

Schöner, G., & Thelen, E. (2006). Using dynamic field theory to rethink infant habituation. Psychological Review, 113(2), 273–299. https://doi.org/10.1037/0033-295X.113.2.273

Scola, C., Holvoet, C., Arciszewski, T., & Picard, D. (2015). Further evidence for infants' preference for prosocial over antisocial behaviors. *Infancy*, 20(6), 684–692. https://doi.org/10.1111/infa.12095

Scott, R. M., Richman, J. C., & Baillargeon, R. (2015). Infants understand deceptive intentions to implant false beliefs about identity: New evidence for early mentalistic reasoning. *Cognitive Psychology*, 82, 32–56. https://doi.org/10.1016/j.cogpsych.2015.08.003

Severson, R. L., & Lemm, K. M. (2016). Kids see human too: adapting an individual differences measure of anthropomorphism for a child sample. Journal of Cognition and Development, 17, 122–141. https://doi.org/10.1080/15248372.2014.989445

Sirois, S., & Jackson, I. (2007). Social cognition in infancy: A critical review of research on higher order abilities. European Journal of Developmental Psychology, 4(1), 46–64.

Slater, A., Von der Schulenburg, C., Brown, E., Badenoch, M., Butterworth, G., Parsons, S., & Samuels, C. (1998). Newborn infants prefer attractive faces. Infant Behavior and Development, 21(2), 345–354.

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

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

Stahl, A. E., & Feigenson, L. (2015). Cognitive development. Observing the unexpected enhances infants' learning and exploration. Science, 348(6230), 91–94. https://doi.org/10.1126/science.aaa3799

Steckler, C. M., Woo, B. M., & Hamlin, J. K. (2017). The limits of early social evaluation: 9-month-olds fail to generate social evaluations of individuals who behave inconsistently. Cognition, 167, 255–265. https://doi.org/10.1016/j.cognition.2017.03.018

Stevens, S. S. (1957). On the psychophysical law. Psychological Review, 64(3), 153-181. https://doi.org/10.1037/h0046162

Sung, J.-Y., Guo, L., Grinter, R.E., & Christensen, H.I. (2007). "My Roomba Is Rambo": Intimate Home Appliances. Proceedings from UbiComp 2007: Ubiquitous

Computing: Lecture Notes in Computer Science, Berlin, Heidelberg. Surian, L., & Franchin, L. (2017). Infants reason about deserving agents: A test with distributive actions. *Cognitive Development*, 44, 49–56. https://doi.org/10.1016/j.

- cogdev.2017.08.009 Tafreshi, D., Thompson, J. J., & Racine, T. P. (2014). An analysis of the conceptual foundations of the infant preferential looking paradigm. *Human Development*, *57*(4), 222–240. https://doi.org/10.1159/000363487
- Thomas, A.J., Saxe, R., & Spelke, E. (2020). Infants use imitation but not comforting or social synchrony to evaluate those in social interactions. In S. Denison, M. Mack, Y. Xu, & B. C. Armstrong (Eds.), Proceedings of the 42nd Annual Meeting of the Cognitive Science Society.
- Thomsen, L., Frankenhuis, W. E., Ingold-Smith, M., & Carey, S. (2011). Big and mighty: preverbal infants mentally represent social dominance. *Science*, 331(6016), 477–480. https://doi.org/10.1126/science.1199198
- Ting, F., Dawkins, M. B., Stavans, M., & Baillargeon, R. (2019). Principles and concepts in early moral cognition. In J. Decety (Ed.), The social brain: A developmental perspective. Cambridge. MA: MIT Press.

Trehub, S. E. (2012). Behavioral methods in infancy: pitfalls of single measures. Annals of the New York Academy of Sciences, 1252, 37–42. https://doi.org/10.1111/j.1749-6632.2012.06448.x

- Tummeltshammer, K. S., Mareschal, D., & Kirkham, N. Z. (2014). Infants' selective attention to reliable visual cues in the presence of salient distractors. *Child Development*, 85(5), 1981–1994. https://doi.org/10.1111/cdev.12239
- Van de Vondervoort, J. W., Aknin, L. B., Kushnir, T., Slevinsky, J., & Hamlin, J. K. (2018). Selectivity in toddlers' behavioral and emotional reactions to prosocial and antisocial others. Developmental Psychology, 54(1), 1–14. https://doi.org/10.1037/dev0000404

Walker-Andrews, A. S. (1997). Infants' perception of expressive behaviors: Differentiation of multimodal information. Psychological Bulletin, 121(3), 437–456. https://doi.org/10.1037/0033-2909.121.3.437

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. Science, 311(5765), 1301–1303. https://doi.org/10.1126/ science.1121448

Warneken, F., & Tomasello, M. (2013). Parental presence and encouragement do not influence helping in young children. Infancy, 18(3), 345–368. https://doi.org/ 10.1111/i.1532-7078.2012.00120.x

Waters, E., Matas, L., & Sroufe, L. A. (1975). Infants' reactions to an approaching stranger: Description, validation, and functional significance of wariness. *Child Development*, 46(2), 348–356. https://doi.org/10.2307/1128127

Waytz, A., Epley, N., & Cacioppo, J. T. (2010). Social cognition unbound: Insights into anthropomorphism and dehumanization. Current Directions in Psychological Science, 19(1), 58–62. https://doi.org/10.1177/0963721409359302

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. Cognition, 69(1), 1-34.

Wynn, K. (1992). Addition and subtraction by human infants. Nature, 358(6389), 749–750.

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., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, 74(1), B1–B11. Yu, C.-L., & Wellman, H.M. (in press), Young Children Treat Puppets (and Dolls and Pictures) Like Real Persons. *Cognitive Development*.

Zhang, F., Jaffe-Dax, S., Wilson, R. C., & Emberson, L. L. (2019). Prediction in infants and adults: A pupillometry study. *Developmental Science*, 22(4), Article e12780.

https://doi.org/10.1111/desc.12780