

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Cognitive Development

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

How (not) to measure infant Theory of Mind: Testing the replicability and validity of four non-verbal measures

Sebastian Dörrenberg^{a,b,*}, Hannes Rakoczy^a, Ulf Liszkowski^b

^a Department of Developmental Psychology, University of Göttingen, Waldweg 26, 37073 Göttingen, Germany

^b Department of Developmental Psychology, University of Hamburg, Von-Melle-Park 5, 20146 Hamburg, Germany

ARTICLE INFO

Keywords:

Infancy
Cognitive development
Implicit Theory of Mind
False belief
Inter-task correlation
Pupil dilation

ABSTRACT

A growing body of infant studies with various implicit, non-verbal measures has suggested that Theory of Mind (ToM) may emerge much earlier than previously assumed. While explicit verbal ToM findings are highly replicable and show convergent validity, systematic replication studies of infant ToM, as well as convergent validations of these measures, are still missing. Here, we report a systematic study of the replicability and convergent validity of implicit ToM tasks using four different measures with 24-month-olds ($N = 66$): Anticipatory looking, looking times and pupil dilation in violation-of-expectation paradigms, and spontaneous communicative interaction. Results of anticipatory looking and interaction-based tasks did not replicate previous findings, suggesting that these tasks do not reliably measure ToM. Looking time and new pupil dilation measures revealed sensitivity to belief-incongruent outcomes which interacted with the presentation order of outcomes, indicating limited evidence for implicit ToM processes under certain conditions. There were no systematic correlations of false belief processing between the tasks, thus failing to provide convergent validity. The present results suggest that the robustness and validity of existing implicit ToM tasks needs to be treated with more caution than previously practiced, and that not all non-verbal tasks and measures are equally suited to tap into implicit ToM processing.

1. Introduction

How does our capacity to understand each other as rational agents with an inner life and subjective perspectives on the world, also known as “Theory of Mind” (ToM), develop? An enormous research program in developmental psychology has been devoted to this question over the last decades (Wellman, 2014). Recently, this research has been revolutionized by new studies with novel methods and surprising findings. In contrast to most tasks traditionally used in ToM research that relied heavily on verbal questions, these new studies have developed completely non-verbal and otherwise simplified, implicit tasks suited for testing even very young infants. The findings from these studies have been received as ground-breaking: They suggest that ToM, in particular the capacity to ascribe false beliefs (FB) to other agents – the litmus test for understanding subjectivity (Wimmer & Perner, 1983) – emerges much earlier than previously assumed in the first months of life (for an overview, see Baillargeon, Scott, & Bian, 2016; Baillargeon, Scott, & He, 2010; Scott & Baillargeon, 2017). A converging line of research suggests that these precocious ToM capacities may remain intact and largely automatic over the lifespan, as indicated by findings that adults often seem to engage in spontaneous yet utterly unconscious ToM processing (Kovács, Téglás, & Endress, 2010; Samson, Apperly, Braithwaite, Andrews, & Scott, 2010; Schneider,

* Corresponding author at: Department of Developmental Psychology, University of Hamburg, Von-Melle-Park 5, 20146 Hamburg, Germany.
E-mail address: sebastian.doerrenberg@uni-hamburg.de (S. Dörrenberg).

<https://doi.org/10.1016/j.cogdev.2018.01.001>

Received 15 February 2017; Received in revised form 15 January 2018; Accepted 16 January 2018

0885-2014/ © 2018 Elsevier Inc. All rights reserved.

Bayliss, Becker, & Dux, 2012; van der Wel, Sebanz, & Knoblich, 2014). Various kinds of such implicit measures have been used with infants, including looking time used with infants as an indicator of violations of expectation (Onishi & Baillargeon, 2005; Surian, Caldi, & Sperber, 2007; Träuble, Marinović, & Pauen, 2010), anticipatory looking (Clements & Perner, 1994; Southgate, Senju, & Csibra, 2007; Surian & Geraci, 2012) and interactive measures such as spontaneous helping (Buttelmann, Carpenter, & Tomasello, 2009; Liszkowski, 2012a, 2012b; ; Southgate, Chevallier, & Csibra, 2010). These studies have produced evidence that in their spontaneous looking and interaction behavior, even very young infants seem capable of engaging in FB representation.

From a theoretical point of view, these findings have been taken as evidence for far-reaching theoretical accounts. According to nativist accounts, the findings suggest that ToM is a domain-specific, probably modular, capacity which is online very early in ontogeny and probably even inborn (e.g., Carruthers, 2013; Leslie, 2005). Standard verbal tasks have failed to uncover these early ToM competencies due to extraneous (linguistic and/or inhibitory) performance factors of the tests. According to recent two-systems accounts, the positive findings from the new implicit tasks reflect an early-developing, evolutionarily more ancient, largely automatic and efficient mindreading system. This system is distinct from and potentially the developmental basis for the later-developing, fully-fledged explicit and flexible ToM system tapped in classical verbal tasks (Apperly & Butterfill, 2009; Low, Apperly, Butterfill, & Rakoczy, 2016).

1.1. Robustness, reliability and replicability of implicit ToM tasks

From an empirical point of view, however, it is still unclear how robust, reliable and replicable these results from the novel implicit measures really are. Questions of reliability and replicability of experimental findings have recently taken center-stage in methodological debates about the evidential status of psychological research (Bakker, van Dijk, & Wicherts, 2012; Button et al., 2013; Makel, Plucker, & Hegarty, 2012; Simmons, Nelson, & Simonsohn, 2011; Simonsohn, Nelson, & Simmons, 2014). In this context, systematic replication attempts across many labs often yield negative results such that existing, often classical, effects cannot be robustly reproduced in independent labs (Open Science Collaboration, 2015). As a consequence, the value and necessity of large-scale and systematic replication studies are now virtually ubiquitously acknowledged in cognitive psychology. In research on automatic ToM in adults, questions of replicability and interpretation of existing results with implicit tasks have recently begun to be addressed (Heyes, 2014; Kovács, Téglás, & Endress, 2016; Phillips et al., 2015; Schneider, Slaughter, & Dux, 2017).

Surprisingly, however, hardly anything is known to date about the robustness, reliability and replicability of implicit ToM findings in infants. This is surprising since reliability issues may be particularly pressing in this area of research: First of all, there are still relatively few established infant studies from implicit measures with positive findings, and most of the published studies have used rather small sample sizes and single trial designs, making them vulnerable to spurious findings (but see Scott & Baillargeon, 2017). Second, to date there are no meta-analyses and we currently do not know about the potential body of unpublished failed replication attempts (the so-called *file-drawer problem*). Third, for most of the published studies there have not been any published replications in independent labs. Fourth, in the few exceptional cases where there are published replication attempts (though they are mostly conceptual, and not direct replications, and often administer multiple within-subject conditions), results are often negative (Grosse Wiesmann, Friederici, Singer, & Steinbeis, 2017; Poulin-Dubois & Yott, 2017; Thoermer, Sodian, Vuori, Perst, & Kristen, 2012; Yott & Poulin-Dubois, 2016; Zmyj, Prinz, & Daum, 2015).

1.2. Convergent validity of implicit ToM tasks

A second fundamental question regarding implicit ToM findings in infants concerns their interpretation and validity. Even if individual implicit ToM tasks turned out to be reliable, this would still not settle issues of validity. What is needed are tests of the convergent validity of individual paradigms. If different tasks are in fact all tapping the same underlying cognitive phenomenon – implicit ToM – then they should converge and correlate. Such correlational patterns of superficially different tasks all designed to tap the same underlying phenomenon have been amply documented for explicit ToM (Astington & Gopnik, 1988; Hamilton, Brindley, & Frith, 2009; Perner & Roessler, 2012; Rakoczy, Bergfeld, Schwarz, & Fiske, 2015). For implicit ToM, however, there hardly have been any analogous studies of convergent validation by correlation. One recent study has investigated diachronic correlations between infant implicit and later explicit ToM measures in a longitudinal design (Thoermer et al., 2012). In this study, an implicit measure (anticipatory looking) in a very specific type of FB task (change-of-location) predicted performance in later explicit FB tasks, but only in superficially analogous (change-of-location) ones and not in other FB tasks. Given the very local nature of this correlation, however, this finding leaves open different interpretations in rich (implicit tasks tap the same kind of ToM processes as later explicit ones) or lean terms (the shared variance between the tasks is reducible to commonalities in the surface features).

With regard to studies of synchronic correlations of various implicit ToM tasks at a given time, to our knowledge there are so far only two studies from one lab. One study (Yott & Poulin-Dubois, 2016) tested infants in a VoE FB task (conceptually after Onishi & Baillargeon, 2005) and in other implicit tasks of their understanding of desires and intentions. Results revealed that – in addition to not replicating the original FB task finding – there was no systematic pattern of inter-task correlations comparable to those found in explicit ToM tasks. Another study (Poulin-Dubois & Yott, 2017) examined 18-month-olds' performances between different ToM constructs. These included a VoE FB task (conceptually after Onishi & Baillargeon, 2005) and an interactive FB task (conceptually after Buttelmann et al., 2009), which both could not be replicated and failed to show any correlations. However, given the diverse, yet un-validated and un-replicated, set of further infant FB tasks, more studies are required that use different tasks to test for convergent validity and the robustness of findings.

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات