This version of the article has been accepted for publication, after peer review but is not the Version of Record and does not reflect post-acceptance improvements or any corrections. The Version of Record is available online at: http://dx.doi.org/10.1007/s11229-024-04809-3

Title: Knowledge First: The Argument from Development
Author: Francesco Antilici francesco.antilici@gmail.com
Abstract: The traditional approach to the analysis of knowledge sees it as a true belief meeting further conditions. I discuss an empirical challenge to this traditional approach, which I call the argument from development. Briefly, the argument is that belief cannot be conceptually prior to knowledge because children acquire the concept of knowledge first. Several prominent scientists and philosophers have argued that this latter claim is supported by many findings with infants and young children. Here, I defend the traditional approach by raising three challenges to the argument from development: the competence-performance challenge, the double-standard challenge, and the underdetermination challenge. I conclude that the developmental data are fully compatible with children acquiring the concept of belief first. In closing, I also argue that further research is needed to investigate when children acquire a concept of knowledge.

1. INTRODUCTION

What is knowledge? The classical view, which some trace back to Plato, is that knowledge is simply justified true belief (JTB). There are well-known counterexamples to this view: cases of true justified beliefs that fall short of knowledge, famously highlighted by Edmund Gettier (1963). In trying to deal with this problem, many have proposed revisions to the JTB analysis (e.g., Clark, 1963), either adding further conditions or replacing the justification condition altogether (e.g., Goldman, 1967). The latter strategy is sometimes presented as a radical departure from the classical view. Even so, both strategies accept the underlying assumption that knowledge is a true belief meeting further conditions. Unfortunately, despite decades of debate and a voluminous literature, there is still no consensus on what those further conditions are. Some have thus started to reconsider the classical view more fundamentally. In his influential book *Knowledge and its Limits*, Timothy Williamson (2000) recommends giving up the project of analysing knowledge in terms of belief or, more generally, in terms of anything more basic. Williamson argues that we can make substantial progress on a variety of epistemological topics by treating knowledge as an "unexplained explainer" (Williamson, 2000, p. 10)

Importantly, although Williamson has focussed on theoretical arguments, he has also acknowledged the relevance of empirical data. In a footnote, for example, Williamson (2000, p. 33, fn. 7) cautiously observes that the traditional approach implies a certain developmental progression: if the concept of belief is a constituent of the concept of knowledge, children should acquire the former before they acquire the latter. Williamson goes on to note that, on the contrary, children seem to acquire the

concept of knowledge first. Call this latter claim *Knowledge-before-Belief,* or *KB*. KB has been recently defended by several authors, including prominent cognitive scientists and philosophers (Nagel, 2013, 2017; Phillips et al., 2020; Phillips & Norby, 2021; Westra & Nagel, 2021; see also Papineau, 2021; Pavese, 2024). These authors (call them *KB-theorists*) argue that empirical evidence converges with Williamson's theoretical arguments, showing that the capacity for knowledge-attribution is more basic than that for belief-attribution in a variety of respects: more evolutionarily ancient, less cognitively demanding, faster, as well as emerging earlier on in development. Phillips et al. (2021, p. 13) close their *BBS* article with a "call to arms", encouraging scientists and philosophers to make knowledge-attribution the focus of their investigations. The knowledge-first revolution is spreading from epistemology to cognitive science.

Here, I focus on whether the developmental data really support KB. I will argue that, when properly understood and contextualised, they do not. After setting the stage in §2, I will raise three challenges to KB-theorists: the competence-performance challenge (§3), the double-standard challenge (§4), and the underdetermination challenge (§5). Together, these challenges cast serious doubt on the argument from development. I then consider whether a specular argument could be made for the claim that belief is conceptually prior to knowledge (§6).

2. SETTING THE STAGE

2.1 The Argument from Development

As we have seen, the argument from development attacks a traditional epistemological assumption, which we can formulate as follows ("TA" stands for Traditional Assumption):

(TA) The concept of belief is a constituent of the concept of knowledge.

The argument further relies on the claim that, if TA is true, then children should acquire the concept of belief before they acquire the concept of knowledge ("BK" stands for Belief-before-Knowledge):

(BK) Children acquire the concept of belief before the concept of knowledge.

We can now formulate the argument as a simple *modus tollens*:

(P1) If TA is the case, then BK is the case.

(P2) BK is not the case.

(C) TA is not the case.

KB-theorists have argued against BK by providing evidence for a mutually exclusive claim, KB:

(KB) Children acquire the concept of knowledge before the concept of belief.

I will argue, instead, that the evidence is fully compatible with BK, undermining P2.

It is worth noting that, as formulated, P1 may also be false. For example, if the concept of belief and the concept of knowledge are both innate, then they may appear at the same time in development, even if the former is in fact a constituent of the latter. Many dismiss nativist positions out of hand, but this attitude does not seem warranted (Margolis & Laurence, 2012). Indeed, as we will see, some evidence suggests that infants may have a concept of belief as early as they have been tested. Nonetheless, I will not pursue this response here, in part because I think that P2 makes for an easier target.

Another issue concerns whether children can be said to have any concepts at all. It is natural to suppose that older children have concepts, but some of the evidence we will discuss concerns pre-verbal infants. The attribution of concepts to non- or pre-linguistic creatures has been defended (see e.g., Carruthers, 2009) but it remains controversial. Still, even if one takes language to be necessary for conceptual thought (Davidson, 1982; McDowell, 1994), it is difficult to explain the relevant behaviours without taking infants to have at least some mental representations, e.g., proto-concepts. In the following, readers may take "concept" as shorthand for "concept or proto-concept". I will also continue to use locutions such as "knowledge-understanding" or "knowledge-attribution", where each should be taken to presuppose possession of the corresponding concept or proto-concept.

Finally, in the developmental literature, "knowledge" is sometimes used in a loose sense, as interchangeable with true belief. The sense of "knowledge" that is relevant here, however, is narrower and more demanding. Since Plato, at least, the philosophical consensus has been that not all true beliefs amount to knowledge. KB-theorists themselves emphasise this distinction (Nagel, 2017, pp. 527–529; Phillips et al., 2021, p. 3); indeed, as we will see, some of their arguments hinge crucially on it. Readers should thus keep in mind that "knowledge" is to be understood in this more demanding sense.

2.2 When Does Belief-Understanding Emerge?

KB-theorists claim that the concept of knowledge is acquired before the concept of belief; but when is the latter acquired? An important piece of evidence comes from the false-belief task (Wimmer & Perner, 1983). In an iconic version, children are told a story about two girls, Sally and Anne. Sally hides her marble in a basket, then goes away; Anne then moves the marble to a box. Children are asked where Sally will look for her marble when she comes back (Baron-Cohen et al., 1985). Most three-year-olds point to the box, as if Sally knew the marble was there; most five-year-olds, in contrast, point to the basket, seemingly understanding that Sally has a false-belief. That this shift happens, albeit with some variability with respect to its timing, is one of the most robust findings in developmental psychology, replicated hundreds of times (see Wellman et al., 2001 for a meta-analysis). And yet, while the finding is established, its interpretation remains controversial. The majority view is that this dramatic shift reflects the acquisition of the concept of belief (Gopnik & Meltzoff, 1997; Gopnik & Wellman, 1994; Perner, 1991; Rakoczy, 2017; Wellman, 2014). Call this the *late-emergence view*. The main alternative is the *early-emergence view*. Early-emergentists maintain that the concept of belief is in place well before children start passing the false-belief task, perhaps already in the first or second year of life. According to early-emergentists, what young children lack is not belief-understanding, but other, auxiliary abilities that are also required to succeed at false-belief tasks (Antilici, 2023; Carruthers, 2013; Helming et al., 2016; Scholl & Leslie, 1999; Setoh et al., 2016; Westra, 2017). There are several findings supporting this view, including studies showing that young children can succeed at false-belief tasks modified to reduce demands on auxiliary abilities (e.g., Setoh et al., 2016; Siegal & Beattie, 1991) or using implicit paradigms (e.g., D. Buttelmann et al., 2009; Onishi & Baillargeon, 2005; Southgate et al., 2007). Some of the relevant evidence will be discussed below.

The debate between late- and early-emergentists has not been resolved yet, and it is not my aim here to resolve it. The important point is that what is or isn't evidence for KB depends in part on whether one takes belief-understanding to emerge "late" or "early". If belief-understanding emerges "early", in the first or second year of life, then KB-theorists must show that knowledge-understanding emerges even earlier, possibly over the first few months of life. If belief-understanding emerges

"late", in the fifth year or after, then KB-theorists have more elbow room; it would be enough to show, for example, that knowledge-understanding emerges in the fourth year. So far, KB-theorists have not committed fully to either path. Instead, they have taken a more pragmatic approach, arguing that the implicit tasks used with infants, as well as the explicit tasks used with older children, reveal the same pattern, with knowledge-understanding emerging first in both cases.

2.3 Evidence from Explicit Tasks

We have seen that most three-year-olds fail the false-belief task (Wellman et al., 2001). KB-theorists follow late-emergentists in taking this as evidence that three-year-olds lack belief-understanding. KB-theorists further argue that four- and five-year-olds fail to attribute justified true beliefs that fall short of knowledge, in so-called Gettier tasks (Phillips et al., 2021, p. 8; Westra & Nagel, 2021, p. 5). In one of the tasks used by Fabricius et al. (2010), for example, Maxi puts his chocolate in the green cupboard, then leaves. While Maxi is away, his sister Anna takes a piece of the chocolate, then considers where to put the rest, and decides to put it back in the green cupboard. Phillips and colleagues argue that, at this point, Maxi has a true belief which is justified yet still "lucky", as in Gettier's (1963) famous cases: it is reasonable for Maxi to expect his chocolate to be where he put it, but in fact his sister Anna moved it and *could* have decided to put it somewhere else (e.g., in the other cupboard). Four-and-a-half-year-olds performed poorly on this task (Fabricius et al., 2010, p. 1406), suggesting they cannot yet attribute mere beliefs. Crucially, KB-theorists argue that at this age children already possess knowledge-understanding (Nagel, 2013, pp. 295–296; Phillips et al., 2021, p. 8). Consider, for example, the knowledge-access task. Children are first shown that a drawer contains a toy dog; then, they are asked whether "Polly", who has "never ever" looked inside the drawer, knows what's inside it. Wellman & Liu (2004, p. 532, Table 4) found that children pass this task a few months before they pass the false-belief task, and follow-up longitudinal studies have confirmed that this holds even when looking at the trajectory followed by individual children (Wellman et al., 2011). Finally, KB-theorists argue that evidence from word learning confirms that the concept of knowledge is acquired first. For example, they claim that children start using "know" before "think", and initially treat "think" as

factive, i.e., they take "S thinks that P" to be true if, and only if, "P" is (Nagel, 2013, pp. 292–293; Phillips et al., 2021, p. 7).

2.4 Evidence from Implicit Tasks

The evidence just discussed comes mostly from explicit tasks, where children are asked questions about the mental state or behaviour of an agent. Given their language demands, explicit tasks are not suited to infants. There are, however, several implicit measures that are more age appropriate, including looking times, spontaneous helping behaviours, or patterns of neural activation. Over the last thirty years or so, research using such implicit measures has revealed that infants are sensitive to other people's mental states, including, among others, knowledge and belief (see Scott et al., 2022 for a review). To illustrate, consider the ground-breaking study by Onishi & Baillargeon (2005). This was a violation-of-expectation study with fifteen-month-olds. The violation-of-expectation paradigm is based on the premise that infants look longer at the event they expect less. In one scenario ("FB-green" condition), Sally puts a toy watermelon slice in a green box on the right. While Sally is away, the watermelon slice moves on its own into the yellow box on the left. Then Sally reappears, and half the infants are shown her reaching for the green box while the other half are shown her reaching for the yellow box. The infants looked longer at the yellow-box event, suggesting they expected Sally to reach for the green box, behaving as if she had a false belief. To see another example, consider the interactive study by D. Buttelmann et al. (2009). In the false-belief condition, when she comes back after the transfer, Sally tries unsuccessfully to open the box where she left her toy (box A), which, unbeknownst to her, has been locked with a pin. If infants understand that Sally is trying to open box A because she thinks her toy is in there, then they should point to or reach for box B, where the toy actually is. This is what D. Buttelmann et al. (2009) found.

Implicit false-belief findings may appear to challenge the late-emergence position, and thus the argument for KB we saw in the previous section. KB-theorists, however, argue that implicit findings overall support KB. First, Phillips et al. (2021, p. 7) claim that there is "uncontroversial evidence" that infants can attribute knowledge already in their first year of life. To illustrate, Phillips and colleagues describe the following violation-of-expectation study with six-month-olds (Luo & Johnson, 2009).

Previous research had shown that, when an agent consistently reaches for one of two objects, infants attribute to her a preference for that object, looking longer when they see her reach for the other one (Woodward, 1998). Luo & Johnson (2009) compared conditions where the agent could see both objects, to conditions where she could only see one (the second object was behind a screen or behind the agent's back). They found that six-month-olds attributed a preference in the former, but not the latter conditions; thus, already at this age, infants appear to consider what the agent does or does not know. In contrast, Phillips et al. (2021, p. 7) claim that there is "comparatively little evidence" of belief-attribution before the first birthday. Most evidence of early belief-attribution, including the findings by Onishi & Baillargeon (2005) and D. Buttelmann et al. (2009) described above, concerns infants already in their second year. Furthermore, Phillips and colleagues cast doubt on these findings by mentioning alternative explanations and failed replications. Consider, for example, the study by Powell et al. (2018). Powell and colleagues made two attempts to replicate the findings by Onishi & Baillargeon (2005), both of which failed; this is one of several failed replications that have surfaced recently, challenging implicit false-belief findings (see Sabbagh & Paulus, 2018 and the other articles in the same issue). Furthermore, while they succeeded in replicating D. Buttelmann and colleagues' finding, Powell and colleagues argue that this is consistent with infants simply attributing knowledge or lack thereof. If Sally does not know where her toy is, she might look in either box; that she tries to open the wrong one is consistent with her ignorance, and it still makes sense for infants to reach for the box containing the toy. Supporting this, Priewasser et al. (2018) found that, if Sally tries to open a third box (box C), most infants still reach for box B. This makes sense if infants take Sally to be ignorant; if they took her to have a false belief, on the other hand, they should have reached for box C. This is because, if Sally thinks that her toy is in box A, and still tries to open box C, she must have some other reason to open that box.

Overall, KB-theorists argue that the evidence that infants attribute knowledge or lack thereof replicates, whereas the evidence that they attribute false beliefs does not. The picture, then, is the following: Knowledge-understanding is already in place by six months or earlier, whereas belief-understanding likely emerges several years later.

3. COMPETENCE V. PERFORMANCE

The case for KB seems compelling at first, but I will argue that it is undermined by three challenges, which I will raise here and in the next two sections. The first challenge we can call the *competence-performance challenge*. There are, in general, two main possible explanations for why participants fail a test for a certain capacity: they may lack the capacity in question (competence failure), or they may have it but struggle to express it, fully or at all (performance failure). The argument for KB assumes that children's difficulties with belief-understanding tasks are of the competence, not performance, kind, but this assumption cannot be had for free; it needs to be defended.

Some readers might not see this as a problem. If there is no evidence whatsoever that young children have belief-understanding, is it not more parsimonious to conclude they have not got it? I rather think that we should suspend judgement; but, more importantly, the antecedent of the conditional is false. There is in fact evidence, to be discussed presently, that children can succeed at the relevant belief-understanding tasks once performance demands are lowered. The competence-performance challenge, then, cannot be ignored or dismissed out of hand; it deserves to be taken seriously.

3.1 False-Belief Tasks

Consider first the finding that three-year-olds fail the explicit false-belief task. KB-theorists follow late-emergentists in inferring that three-year-olds lack belief-understanding. As mentioned above (§2.2), however, early-emergentists defend a performance explanation for this finding: they argue that three-year-olds already possess belief-understanding but fail to express it because they lack certain auxiliary abilities that are also required for succeeding at false-belief tasks, including executive-functioning, language-ability, pragmatic-understanding, and others. Importantly, many studies have found that, when demands on these auxiliary abilities are reduced, children succeed at a younger age (e.g., Hansen, 2010; Rubio-Fernández & Geurts, 2013; Salter & Breheny, 2019; Setoh et al., 2016; Siegal & Beattie, 1991). Siegal & Beattie (1991), for example, suggest that children, due to their poor pragmatic-understanding, may take the experimenter to be asking where Sally will have to look for the marble to find it. To test this hypothesis, they added a temporal marker to the question: Where will Sally look *first* for her marble? Lo and behold, most three-year-olds now succeeded, with their

performance on a true-belief control unaffected (Siegal & Beattie, 1991, pp. 6–7; see Surian & Leslie, 1999 for a replication). To see another example, Setoh et al. (2016), using a low-inhibition false-belief task (where the marble is removed from the scene) with two response-generation practice-trials (e.g., "Where is the ball?", when shown a picture of a ball and a Frisby), obtained success with even younger children (two-and-a-half-year-olds), a finding which has been replicated (Grosso et al., 2019; Kaltefleiter et al., 2021; Scott et al., 2020; Sodian et al., 2024).

3.2 Gettier Tasks

Consider now the finding that four- to six-year-olds fail the Gettier task (Fabricius et al., 2010). Fabricius and colleagues actually found a U-shaped pattern: three- and six-year-olds performed above-chance, while four- and, on some tasks, five-year-olds, dipped at chance (Fabricius et al., 2010, p. 1406). Other studies have found an even deeper and wider trough, with three- and ten-year-olds above-chance but four-, five- and six-year-olds below-chance, and eight-year-olds at chance (Oktay-Gür & Rakoczy, 2017, p. 35). Notably, this U-shaped pattern challenges not only early-emergence accounts, but also the late-emergence orthodoxy, suggesting that belief-understanding may emerge even later than late-emergentists have traditionally assumed. When they have discussed this finding, late-emergentists have typically resisted revising their position, proposing performance explanations instead (Huemer et al., 2023; Oktay-Gür & Rakoczy, 2017; Rakoczy & Oktay-Gür, 2020; Schidelko et al., 2022). To illustrate, Oktay-Gür & Rakoczy (2017) suggest that older children may find the true-belief question just too trivial, thus taking the experimenter to be asking something else. To test this hypothesis, they modified the scenario by adding a second agent, who does not witness the transfer. The contrast between the agent with a true-belief and the one with a false-belief makes the true-belief question less trivial. In this context, four- and six-year-olds performed above chance (Oktay-Gür & Rakoczy, 2017, p. 37).

3.3 Word Learning

The competence-performance challenge comes up when discussing the evidence from word-learning as well. Recall that KB-theorists claim that children start using "know" before "think", and that when

they start using "think", they initially treat it as a factive verb, judging sentences of the form "S thinks that P" to be true if and only if P is true (§2.3). Whether children's early uses of "know" really are references to knowledge, as opposed to mere conversational uses, is debated (see Harris et al., 2017; Shatz et al., 1983 for opposing positions). In any case, even if children do start talking about knowledge before they talk about belief, this and the fact that they initially treat "think" as factive (e.g., de Villiers & Pyers, 2002) may be explained by the difficulty of understanding that "think" expresses the concept of belief. Note that the pragmatics of "think" can be very confusing (Westra, 2017). For example, uttering "S thinks P" can be a way to indirectly assert that P. Lewis et al. (2017) suggest that young children may struggle differentiating such indirect uses from genuine belief-reports. If, when the experimenter utters "S thinks that P", children take her to be asserting that P, this may explain why they judge the utterance to be true if and only if P is true. To test their pragmatic hypothesis, Lewis et al. (2017) introduced a second character with a different belief. The thought was that the conflict between the two beliefs could help children recognize that the sentence they were asked to judge was a belief-report. The addition of the second character brought three-year-olds from below-chance to at-chance (Lewis et al., 2017, p. 11), suggesting that pragmatics does indeed play a role.

4. Double Standards

Overall, the findings discussed in §3 suggest that young children's difficulties with belief-understanding tasks may be of the performance, not competence variety, undermining an important argument for KB. Importantly, my claim is not that this evidence is conclusive, but that it deserves to be taken seriously and addressed, something which KB-theorists have not yet done. And I will now argue that, even if performance explanations turn out to be incorrect, the argument for KB faces another important challenge, the *double-standard challenge*.

4.1 Explicit Tasks

KB-theorists have followed late-emergentists in taking failure on belief-understanding tasks to be evidence that young children do not yet have a concept of belief. When we look at how late-emergentists

motivate this conclusion, however, we see that their argument supports a broader claim, namely that young children do not have adult-like concepts of any representational states, including belief *and knowledge*. The double-standard challenge arises because KB-theorists endorse the late-emergence argument so far as it concerns belief-understanding but seem to overlook its implications concerning knowledge-understanding.

Late-emergentists have generally operated within a theory-theory framework, according to which children develop, test, and revise their own theories as they accumulate more evidence about the relevant domain (Gopnik & Wellman, 1994; Perner, 1991; Rakoczy, 2017). A common assumption within this framework is that there is a constitutive link between concepts and theories. To have the same concepts adults have, children must have the same theories. Now, a distinctive feature of the adult theory is that it posits representational mental states. The hallmark of representation is, of course, the possibility of misrepresentation: when one represents something, one may represent it falsely or inaccurately. The aim of the false-belief task is to test whether children's theory of mind is a representational one, in this sense. Thus, roughly, the reasoning is as follows: since three-year-olds fail the false-belief task, they do not have a representational theory of mind; and without a representational theory of mind, they cannot have any of the concepts embedded in that theory, including, among others, the concept of belief. Now, the problem is that knowledge is every bit as representational as belief. Both figure in the adult theory as representational states. A child who does not have that theory cannot have either concept for the very same reason. Three-year-olds may have *some* understanding of knowledge and belief but, the story goes, this can only be a very limited understanding, not quite the representational, adult-like one older children possess.

The same problem affects KB-theorists' interpretation of Gettier true-belief findings. Fabricius and colleagues, for example, maintain that it is only when they start passing both false-belief and Gettier tasks that children can be said to have acquired a representational theory of mind (see Fabricius et al., 2021, e.g., p. 17). Once again, this entails that children who fail either task cannot have concepts of knowledge or belief. This aspect of Fabricius and colleagues' view can be misunderstood, because

they argue that children who pass the false- belief task but fail the Gettier task rely on generalisations like the following:

(1) sees → knows → gets it right

(2) doesn't see → doesn't know → gets it wrong

Note that these generalisations are phrased in terms of "knowledge". For Fabricius and colleagues, however, children in this transitional phase have a rather limited understanding of knowledge, as a nonrepresentational state that can only arise from direct perceptual access (as opposed to, say, inference or memory) and which directly causes successful (or, in its absence, unsuccessful) actions (Fabricius et al., 2021, p. 18). The representational, adult-like concept is only acquired once children start passing both false-belief and Gettier tasks.

Finally, this problem also affects KB-theorists interpretation of word-learning data. As we have seen, children start using "know" and "think" in their third year, well before they start passing knowledge-access, false-belief and Gettier tasks. Since at that age children (allegedly) lack a representational theory of mind, their use of "know" and "think" cannot yet express the corresponding concepts.

4.2 Rejecting the Orthodoxy

Perhaps unsurprisingly, some KB-theorists have rejected some of the assumptions typically made by late-emergentists. For example, Nagel (2017, p. 535) takes issue with the assumption that, to have a concept of knowledge, one must be aware that one may represent the world inaccurately. Instead, Nagel suggests that it would be enough for one to be aware that different people can represent the world differently. Crucially, two people can represent the world differently without either representing it inaccurately. If I take you not to know something, I take your representation of the world to be different from my own, but this does not make your representation inaccurate; just incomplete. By this standard, a child who fails the false-belief task may still count as having a concept of knowledge. More radically, Phillips & Norby (2021, pp. 8–9) argue that to represent something, it is enough –

roughly[1] – that one has an ability to track it; that is, to respond differently to its presence and absence in many situations. If one responds differently to situations where one knows v. doesn't know, one can be said to represent knowledge. This is true even if one lacks a "sophisticated understanding" of its nature, of the kind embodied in the adult theory. Once again, the conclusion is that children who pass the knowledge-access task but fail the false-belief task can still be credited with a concept of knowledge.

However, even if Nagel (2017) and Phillips & Norby (2021) are right to lower the standards for knowledge-understanding this way, this is not enough to fix the problem. For we will then have to lower them similarly for belief-understanding; at a minimum, the use of different standards would require justification. Note that three-year-olds do pass some belief-understanding tasks, such as the diverse-belief task (previously known as the discrepant-belief task, see Wellman & Bartsch, 1988). The diverse-belief task is meant to test whether children can attribute beliefs that differ from their own, even though they do not know which is true and which is false. Participants are told that Linda is looking for her cat, which could be in either of two locations, and asked to guess which; then, they are told that Linda believes it to be in the other location and asked where she will look for it (Wellman & Liu, 2004)[2]. Importantly, the diverse-belief task is passed before the false-belief task and, in Western children albeit not Chinese or Iranian ones, even before the knowledge-access task (Shahaeian et al., 2011; Wellman et al., 2011). One may wonder why passing the diverse-belief task is not enough to demonstrate belief-understanding. What is missing, according to Wellman, is precisely the understanding that others may represent the world *inaccurately*. But note that the task does show that three-year-olds can track beliefs in many situations and understand that others may represent the world

[1] Phillips & Norby (2021, fn. 3) acknowledge that most tracking theories have further necessary conditions. For their argument to work, however, it must be possible to satisfy those further conditions without having a sophisticated understanding of whatever one is tracking.

[2] Westra & Carruthers (2017) suggest that three-year-olds may take utterances like "Linda thinks the cat is in the garden" to imply that that's where the cat is; if they are right, then success on the diverse-belief task may not evidence belief-understanding. However, Westra & Carruthers (2017) also argue that three-year-olds fail the false-belief task for similar reasons (e.g., they take a question like "Where does Sally think her toy is?" to concern the actual location of the toy). If they are right, the competence-performance challenge discussed in §3 presents a more urgent concern than the double-standard challenge. (I thank an anonymous reviewer for encouraging me to clarify this point).

differently from them. Arguably, then, by Nagel's and Phillips & Norby's lower standards, three-year-olds do have a concept of belief.

Overall, with the higher standards late-emergentists have traditionally employed, three-year-olds understand neither knowledge nor belief; with the lower standards KB-theorists have proposed, they understand both; but in neither case they have knowledge-understanding without belief-understanding.

4.3 Implicit Tasks

The same issue comes up when interpreting implicit false-belief findings with infants. Phillips et al. (2021, p. 7) cast doubt on these findings noting that they are open to alternative, "sceptical" interpretations; they cite specifically Burge (2018), Butterfill & Apperly (2013), Heyes (2014), Priewasser et al. (2018) and Wellman (2014). Whether any of these sceptical proposals can really explain implicit findings is a matter of controversy (see e.g., Carruthers, 2013; Scott et al., 2022), and one I cannot resolve here. My point, though, is that on these sceptical interpretations, infants understand neither belief *nor* knowledge. This is not surprising: what drives the scepticism is precisely the conviction that infants are unlikely to have a representational theory of mind, and without one – the argument goes – they can understand neither knowledge nor belief.

To illustrate, some of the sceptics cited above deny that infants attribute any mental states at all. Heyes (2014) argues that infants habituate to low-level features of the visual stimuli, such as shape and colour. Burge (2018) argues that infants predict action by attributing nonmental sensory and conative states. Priewasser et al. (2018) suggest that infants may simply expect agents to pursue desirable goals. Others concede that infants attribute some mental states but argue that those mental states are substantially simpler than the propositional attitudes figuring in the adult theory. Butterfill & Apperly (2013), for example, argue that infants attribute mental states called "encountering" and "registration", which behave a bit like perception and belief, while being simpler and nonrepresentational. Wellman (2014) argues that infants attribute states of "knowledge" or lack thereof, which may at first seem in line with KB. However, recall that Wellman has been one of the main promoters of the theory

that children only acquire a representational theory of mind around the age of four. More carefully stated, his position is that infants simply *track knowledge or knowledge-like states*:

> For simplicity, I will refer to this infant understanding of engagement as tracking knowledge or knowledge-like experiences. If something happens and the agent is appropriately engaged with it, she is aware of it and in that limited sense knows of it; if it happened and she was unaware, she does not know. I do not think such young children track knowledge exactly as adults more fully understand knowing. (Wellman, 2014, p. 178)

Once again, one can argue that tracking knowledge is enough to represent it (Phillips & Norby, 2021). But the problem is that this also removes the main reason for being sceptical that infants have belief-understanding. Many of the sceptics do in fact concede that infants show sensitivity to belief-involving situations. Butterfill & Apperly (2013), for example, explicitly acknowledge that infants can track beliefs, both true and false; yet they deny that infants represent beliefs as such precisely because, they argue, tracking something is not enough to represent it.

5. UNDERDETERMINATION

Together, the competence-performance and double-standard problems pose a formidable challenge to KB-theorists. If the findings KB-theorists appeal to are open to performance explanations, they do not support KB; but even if they are not, the motivation offered for taking them as evidence that young children lack belief-understanding would entail that they lack knowledge-understanding as well. Either way, the argument for KB is problematic.

I will now highlight a third problem, the *underdetermination* challenge, which provides context for the other two. This isn't the general underdetermination problem raised by Duhem and Quine (Kyle, 2023), but a more specific issue that arises in differentiating experimentally between knowledge-attribution and belief-attribution. The problem arises because, even if knowledge is not a kind of belief, there is still substantial overlap between the two: most, perhaps all, situations where one knows something are also situations where one believes it, and many situations where one doesn't know something are also situations where one has no beliefs on the matter. Furthermore, knowledge and belief often support the same behavioural predictions. Regardless of whether Sally knows that the marble is in the basket, or merely believes it, she will look for it there. When using scenarios where knowledge and belief overlap, it is thus difficult to tell whether children are attributing knowledge or

belief. Potentially, the problem can be overcome using scenarios where knowledge and belief do not overlap; for example, situations where an agent's belief is false, true but unjustified, or true-and-justified but nonetheless lucky, as in Gettier cases. If children attribute knowledge but not belief, they should treat all these situations as ignorance situations. The question is whether there are any data supporting this prediction, and whether they can escape the competence-performance and double-standard challenges.

5.1 Explicit Tasks

Consider the knowledge-access task, described above (§2.3). Children are asked whether Polly, who has "never ever" looked inside the drawer, knows what's inside it. Most three-year-olds answer "no", and KB-theorists take this as evidence that they possess knowledge-understanding. However, note that in this case knowledge and belief overlap. The agent doesn't know what is inside the drawer, nor does she have any beliefs on the matter. It is true that children are asked whether the agent *knows*, but for all we know, children may take "know" to mean "believe" or "believe truly". Thus, the underdetermination problem: the evidence is consistent with the claim that children attribute knowledge, but it is also consistent with the claim that they attribute belief, so it cannot be used to show that they have the former capacity but not the latter.

KB-theorists may object that, at an age where they pass the knowledge-access task, children fail Gettier tasks, showing they can attribute knowledge but not beliefs. We have already seen some problems with KB-theorists' interpretation of Gettier tasks (§3.2, §4.1), and when examining the finding more closely, a further problem becomes apparent. Although they used a Gettier task, as described in §2.3, Fabricius et al. (2010) also used other true-belief tasks where the agent arguably knows (as opposed to merely having a justified true belief about) the relevant facts. For example, in one task (Study 2, modified true-belief location task), the object is put in one location, then transferred to a second location; the agent watches this transfer, then leaves and comes back, with nothing happening in her absence. Crucially, five-year-olds did not perform any better on this task compared to the Gettier task (Fabricius et al., 2010, p. 1410). The other references Phillips et al. (2021) cite as "Gettier" findings (Oktay-Gür & Rakoczy, 2017; Perner et al., 2015) did not in fact use any Gettier tasks, but

only "normal", knowledge/true-belief tasks. *Pace* Phillips et al. (2021), then, the finding isn't that children attribute knowledge but not true belief; it is that they struggle to attribute either, and *this* finding does not help solve the underdetermination problem.

KB-theorists may further object that when they start passing the knowledge-access task, children still fail the false-belief task, showing they lack belief-understanding. But first, we have seen that three-year-olds' difficulties with the false-belief task may be of the performance, not competence, kind (§3.1). Furthermore, we have seen that the argument offered for taking failure on the false-belief task as evidence that young children lack belief-understanding would entail that they lack knowledge-understanding as well (§4.1). Either way, it is difficult for KB-theorists to rely on negative false-belief findings to overcome the underdetermination challenge.

5.2 Implicit Tasks

The underdetermination challenge similarly undermines the evidence from implicit tasks. Consider the study by Luo & Johnson (2009), described above (§2.4). The study compared a knowledge-condition, where the agent can see that there are two objects on the table, to an ignorance-condition, where the agent can only see one object. The problem is that in the knowledge-condition, the agent both knows and believes there to be two objects, whereas in the ignorance-condition, the agent neither knows nor believes there to be a second object. In these situations, then, knowledge and belief overlap. The finding that infants only attribute a preference in the knowledge-condition is consistent with the hypothesis that they attribute knowledge (or lack thereof), but it is also consistent with the hypothesis that they attribute belief (or lack thereof); hence the underdetermination problem.

Recall that Phillips et al. (2021, p. 7) claim there is "comparatively little evidence" of belief attribution in the first year; what they must mean is that there is comparatively little evidence of *false-belief* attribution in the first year. In any case, there is more evidence for this than Phillips et al. acknowledge (Hirshkowitz & Rutherford, 2021; Hyde et al., 2018; Kampis et al., 2015; Kovács et al., 2010; Luo, 2011; Southgate & Vernetti, 2014). For example, using a set-up similar to that of Luo & Johnson (2009), Luo (2011) found that ten-month-old infants take an agent reaching for an object as expression of a preference if the agent falsely believes there to be a second object behind a screen.

Importantly, infants demonstrate belief-understanding as early as they have been tested. The study by Southgate & Vernetti (2014), for example, found that six-month-olds anticipate an agent reaching for a box when the agent falsely believes that the box contains a ball, but not when the agent falsely believes the box does not contain a ball.

As mentioned above (§2.4), some findings suggest that, in some implicit false-belief studies (e.g., D. Buttelmann et al., 2009) infants succeed by attributing knowledge or lack thereof instead of belief (Powell et al., 2018; Priewasser et al., 2018). However, other false-belief tasks are not open to this interpretation (e.g., Luo, 2011; Onishi & Baillargeon, 2005), and some studies have explicitly compared ignorance and false-belief conditions, finding different results in each (He et al., 2011; Knudsen & Liszkowski, 2012; Scott et al., 2010; Scott & Baillargeon, 2009). For example, Knudsen & Liszkowski (2012) found that infants used pointing gestures to warn an agent about yucky stuff in containers; however, when the agent falsely believed the toy she was looking for to be in a specific container, the infants pointed significantly more often to that container, whereas when the agent didn't know where her toy was, they pointed to both containers.

Of course, implicit false-belief findings have been very controversial; worries have been raised about their robustness and validity (see Baillargeon et al., 2018; Poulin-Dubois et al., 2018 for discussion). Some readers may thus not feel persuaded by this evidence. I do not mean to downplay such worries, nor can I resolve the controversy here. Arguably, however, the evidence that infants can attribute knowledge and ignorance raises similar concerns. Indeed, as noted in §4.3, many have argued that infants can succeed at implicit tasks without understanding either knowledge or belief; and I will argue below that the failed replications similarly play no favourites.

5.3 Failed Replications

Many are currently sceptical about implicit findings due to the failed replications. KB-theorists claim that some of those findings, namely those consistent with early knowledge-attribution, have replicated, whereas others, namely those suggestive of false-belief attribution, have not (§2.4). If correct, this would support KB. A cursory look at the data suggests a more complex pattern, however. Powell et al. (2018) succeeded in replicating D. Buttelmann et al.'s (2009) results in the false-belief

condition, which they argue is really an ignorance condition. The problem is that other attempts to replicate that same finding have failed (Crivello & Poulin-Dubois, 2018; Poulin-Dubois & Yott, 2018). Even Priewasser et al. (2018) report mixed findings in their first study, replicating D. Buttelmann et al.'s (2009) result in the false-belief condition but not in the true-belief (or "knowledge") condition. It is not true, then, that the evidence consistent with knowledge-attribution has straightforwardly replicated; some has not.

On the other hand, while Powell et al. (2018) could not replicate Onishi & Baillargeon's (2005) finding, this is not enough to dismiss it. A meta-analysis by Barone et al. (2019) found that, despite the failed replications, infants are still more likely to succeed at violation-of-expectation false-belief tasks than to fail, and more positive findings have been published since (F. Buttelmann & Kovács, 2019; Hirshkowitz & Rutherford, 2021; Schulze & Buttelmann, 2022; Tauzin & Gergely, 2018). There remain worries concerning heterogeneity and publication bias (see Barone et al., 2019), but it is nonetheless incorrect to say that violation-of-expectation false-belief findings have not replicated; the results are mixed, but arguably more positive than negative[3].

Overall, the evidence is more complex than suggested by some KB-theorists, and much less supportive of their view. Just as there is evidence that young infants attribute knowledge and ignorance, there is also evidence that they attribute false beliefs (§5.2). Many have raised doubts concerning the robustness and validity of the latter, but the former is just as open to sceptical interpretations (§4.3) and similarly challenged by recalcitrant findings, including some failed replications of knowledge and ignorance conditions on the one hand, and some successful replications of false-belief conditions on the other. It is too early to tell how the replication crisis will play out, and I do not mean to prejudge the issue one way or the other. My point is that, currently, the same type of worries that have been raised about the evidence of early false-belief attribution can also be raised about the

[3] What of Holland & Phillips (2020) meta-analysis, which found that "factive" (i.e., knowledge) conditions are more replicable than "nonfactive" (i.e., false-belief) conditions? This meta-analysis has important limitations: it considered only a fraction of the evidence and included studies with adults and older children in addition to studies with infants (only the latter are relevant here). Crucially, two-thirds of the "factive" tests came from a study with adults, with the two studies contributing the remaining third showing no difference in replicability between factive and nonfactive tests (see Holland & Phillips, 2020, Table 1). (I thank an anonymous reviewer for directing me to this important resource.)

evidence of early knowledge attribution. Accordingly, it is difficult to dismiss the former without undermining the latter.

6. An Argument for TA?

I have argued that the developmental evidence does not support KB over BK[4]. However, one may wonder whether a stronger conclusion may be warranted; that is, whether we should take the evidence to positively support BK. If we could establish BK empirically, we could then build an explanatory argument for TA, and thus against Williamson's Knowledge-First approach. If children acquire the concept of belief first (BK), this may then be explained, and perhaps *best* explained, by taking the concept of belief to be a constituent of the concept of knowledge (TA).

Does the evidence, then, support BK? We have seen that the emergence of belief-understanding remains controversial; but even harder, at this point, is to locate the emergence of knowledge-understanding. One of the upshots of the underdetermination problem is that all the evidence that children attribute knowledge or lack thereof can also be explained by taking them to attribute belief or lack thereof instead. Recall the Gettier and knowledge/true-belief tasks used by Fabricius et al. (2010). We have seen that children perform similarly on the two tasks (§5.1). Crucially, when children succeed, they may succeed by attributing justified true beliefs alone. After all, in both Gettier and knowledge/true-belief tasks the protagonist has a justified true belief, and the same behavioural prediction follows: Whether Sally knows, or merely has a justified true belief, that the marble is in the basket, she will look for it there. Since all the evidence we have discussed can be explained in terms of belief-attribution, it is unclear when knowledge-understanding emerges. This is an obstacle to establishing KB, but it is also an obstacle to establishing BK.

[4] As noted in the introduction, KB is part of a broader theoretical perspective where knowledge-attribution is seen as more evolutionarily ancient and less cognitively-demanding than belief-attribution. Perhaps, such considerations could be leveraged to argue that KB is a better explanation than BK. This is a question that goes beyond the scope of the present article, and which I leave for further research to investigate. Nonetheless, KB is a developmental claim, so the developmental data are the most relevant to evaluate it. (I thank an anonymous reviewer for raising this point.)

What is needed to show that children have a concept of knowledge is a situation where the agent would behave differently depending on whether she has knowledge or mere justified true belief. Williamson (2000, Chapters 2–3) argues that knowledge plays an ineliminable role in psychological explanation, one which cannot be played by mere justified true beliefs (or indeed, in his eyes, by any belief-based composite condition). Specifically, knowledge is more resistant to counterevidence than mere belief, and thus more likely to support persistent action. If a burglar merely believes, as opposed to knowing, that there is a diamond in the house she is ransacking, she is more likely to encounter evidence that defeats her belief, and thus to give up her search (Williamson, 2000, pp. 62–63). If this is on the right track, one could adapt one of Williamson's scenarios to investigate whether children expect knowledgeable actors to behave differently than those who merely have a justified true belief. However, the claim that knowledge plays an ineliminable role in psychological explanation remains controversial (see e.g., Kipper, 2018; Magnus & Cohen, 2003).

In the absence of evidence that children attribute knowledge, some may argue that we should go with the simplest hypothesis. Since we have independent reason to take children to have a concept of belief (e.g., their performance on false-belief tasks), it may be more parsimonious to take them to have only that concept, as opposed to concepts of both belief *and* knowledge. But I am reluctant to put this much weight on parsimony, and I hope that future research will shed more light on the emergence of knowledge-understanding. Thus, although I have disagreed with the empirical claims made by Phillips et al. (2021), I endorse their call to arms. More research is needed to figure out when the concept of knowledge is acquired, and this may well broaden our understanding of the nature of knowledge.

REFERENCES

Antilici, F. (2023). A new rationalist account of the development of false-belief understanding. *Philosophical Studies*, *180*(9), 2847–2870.

Baillargeon, R., Buttelmann, D., & Southgate, V. (2018). Invited Commentary: Interpreting failed replications of early false-belief findings: Methodological and theoretical considerations. *Cognitive Development*, *46*(June), 112–124.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, *21*(1), 37–46.

Barone, P., Corradi, G., & Gomila, A. (2019). Infants' performance in spontaneous-response false belief tasks: A review and meta-analysis. *Infant Behavior and Development*, *57*(February), 101350.

Burge, T. (2018). Do infants and non-human animals attribute mental states? *Psychological Review*, *125*(3), 409–434. https://doi.org/10.1037/rev0000091

Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-Month-Old Infants Show False Belief Understanding in an Active Helping Paradigm. *Cognition*, *112*(2), 337–342.

Buttelmann, F., & Kovács, Á. M. (2019). 14-Month-olds anticipate others' actions based on their belief about an object's identity. *Infancy*, *24*(5), 738–751. https://doi.org/10.1111/infa.12303

Butterfill, S. A., & Apperly, I. A. (2013). How to Construct a Minimal Theory of Mind. *Mind & Language*, *28*(5), 606–637.

Carruthers, P. (2009). Invertebrate concepts confront the generality constraint (and win). In R. W. Lurz (Ed.), *The Philosophy of Animal Minds* (pp. 89–107). Cambridge University Press.

Carruthers, P. (2013). Mindreading in infancy. *Mind and Language*, *28*(2), 141–172.

Clark, M. (1963). Knowledge and Grounds: A Comment on Mr. Gettier's Paper. *Analysis*, *24*(2), 46–48.

Crivello, C., & Poulin-Dubois, D. (2018). Infants' false belief understanding: A non-replication of the helping task. *Cognitive Development*, *46*, 51–57.

Davidson, D. (1982). Rational Animals. *Dialectica*, *36*(4), 317–327.

de Villiers, J. G., & Pyers, J. E. (2002). Complements to Cognition: a Longitudinal Study of the Relationship Between Complex Syntax and False-Belief-Understanding. *Cognitive Development*, *17*, 1037–1060.

Fabricius, W. V., Boyer, T. W., Weimer, A. A., & Carroll, K. (2010). True or False: Do 5-Year-Olds Understand Belief? *Developmental Psychology*, *46*(6), 1402–1416.

Fabricius, W. V., Gonzales, C. R., Pesch, A., Weimer, A. A., Pugliese, J., Carroll, K., Bolnick, R. R., Kupfer, A. S., Eisenberg, N., & Spinrad, T. L. (2021). Perceptual Access Reasoning (PAR) in

Developing a Representational Theory of Mind. *Monographs of the Society for Research in Child Development*, *86*(3), 7–154. https://doi.org/10.1111/mono.12432

Gettier, E. (1963). Is Justified True Belief Knowledge ? *Analysis*, *23*(6), 121–123.

Goldman, A. I. (1967). A Causal Theory of Knowing. *The Journal of Philosophy*, *64*(12), 357–372.

Gopnik, A., & Meltzoff, A. N. (1997). *Words, Thoughts, and Theories (Learning, Development, and Conceptual Change)*. MIT Press.

Gopnik, A., & Wellman, H. M. (1994). The Theory Theory. In L. A. Hirschfield & S. A. Gelman (Eds.), *Mapping the Mind: Domain Specificity in Cognition and Culture* (pp. 257–293). Cambridge University Press.

Grosso, S. S., Schuwerk, T., Kaltefleiter, L. J., & Sodian, B. (2019). 33-month-old children succeed in a false belief task with reduced processing demands: A replication of Setoh et al. (2016). *Infant Behavior and Development*, *54*(August 2018), 151–155.

Hansen, M. B. (2010). If you know something, say something: Young children's problem with false beliefs. *Frontiers in Psychology*, *1*(23), 1–7.

Harris, P. L., Yang, B., & Cui, Y. (2017). 'I Don't Know': Children's Early Talk About Knowledge. *Mind and Language*, *32*(3), 283–307. https://doi.org/10.1111/mila.12143

He, Z., Bolz, M., & Baillargeon, R. (2011). False-Belief Understanding in 2.5-Year-Olds: Evidence from Violation-Of-Expectation Change-Of-Location and Unexpected-Contents Tasks. *Developmental Science*, *14*(2), 292–305.

Helming, K. A., Strickland, B., & Jacob, P. (2016). Solving the Puzzle About Early Belief-Ascription. *Mind & Language*, *31*(4), 438–469.

Heyes, C. M. (2014). False Belief in Infancy: a Fresh Look. *Developmental Science*, *17*(5), 647–659.

Hirshkowitz, A., & Rutherford, M. D. (2021). Longer looking to agent with false belief at 7 but not 6 months of age. *Infant and Child Development*, *30*(5). https://doi.org/10.1002/icd.2263

Holland, C., & Phillips, J. (2020). A theoretically driven meta-analysis of implicit theory of mind studies: The role of factivity. *Proceedings of the Annual Meeting of The Cognitive Science Society*, *42*.

Huemer, M., Schröder, L. M., Leikard, S. J., Gruber, S., Mangstl, A., & Perner, J. (2023). The knowledge ("true belief") error in 4- to 6-year-old children: When are agents aware of what they have in view? *Cognition*, *230*. https://doi.org/10.1016/j.cognition.2022.105255

Hyde, D. C., Simon, C. E., Ting, F., & Nikolaeva, J. (2018). Functional organization of the temporal-parietal junction for theory of mind in preverbal infants: A near-infrared spectroscopy study. *The Journal of Neuroscience*, *38*(18), 0264–17.

Kaltefleiter, L. J., Sodian, B., Kristen-Antonow, S., Grosse Wiesmann, C., & Schuwerk, T. (2021). Does syntax play a role in Theory of Mind development before the age of 3 years? *Infant Behavior and Development*, *64*(May), 101575. https://doi.org/10.1016/j.infbeh.2021.101575

Kampis, D., Parise, E., Csibra, G., & Kovács, Á. M. (2015). Neural Signatures for Sustaining Object Representations Attributed to Others in Preverbal Human Infants. *Proceedings of the Royal Society B*, *282*(20151683).

Kipper, J. (2018). Acting on true belief. *Philosophical Studies*, *175*(9), 2221–2237. https://doi.org/10.1007/s11098-017-0956-8

Knudsen, B., & Liszkowski, U. (2012). 18-Month-Olds Predict Specific Action Mistakes Through Attribution of False Belief, Not Ignorance, and Intervene Accordingly. *Infancy*, *17*(6), 672–691.

Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The Social Sense: Susceptibility to Others' Beliefs in Human Infants and Adults. *Science*, *330*(6012), 1830–1834.

Kyle, S. (2023). Underdetermination of Scientific Theory. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Summer 2023). URL = <https://plato.stanford.edu/archives/sum2023/entries/scientific-underdetermination/>.

Lewis, S., Hacquard, V., & Lidz, J. (2017). "Think" Pragmatically: Children's Interpretation of Belief Reports. *Language Learning and Development*, *13*(4), 395–417. https://doi.org/10.1080/15475441.2017.1296768

Luo, Y. (2011). Do 10-Month-Old Infants Understand Others' False Beliefs? *Cognition*, *121*(3), 289–298.

Luo, Y., & Johnson, S. C. (2009). Recognizing the Role of Perception in Action at 6 Months. *Developmental Science*, *12*(1), 142–149.

Magnus, P. D., & Cohen, J. (2003). Williamson on Knowledge and Psychological Explanation. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *116*(1), 37–52.

Margolis, E., & Laurence, S. (2012). In Defense of Nativism. *Philosophical Studies*, *165*(2), 693–718.

McDowell, J. (1994). *Mind and World*. Harvard University Press.

Nagel, J. (2013). Knowledge as a Mental State. In T. S. Gendler & J. Hawthorne (Eds.), *Oxford Studies in Epistemology* (Vol. 4, pp. 272–308). Oxford University Press. https://doi.org/10.1093/acprof

Nagel, J. (2017). Factive and nonfactive mental state attribution. *Mind and Language*, *32*(5), 525–544. https://doi.org/10.1111/mila.12157

Oktay-Gür, N., & Rakoczy, H. (2017). Children's difficulty with true belief tasks: Competence deficit or performance problem? *Cognition*, *166*, 28–41.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-Month-Old Infants Understand False Beliefs? *Science*, *308*(5719), 255–258.

Papineau, D. (2021). The disvalue of knowledge. *Synthese*, *198*(6), 5311–5332. https://doi.org/10.1007/s11229-019-02405-4

Pavese, C. (2024). Factive Mindreading in the Folk Psychology of Action. In A. Logins & J.-H. Vollet (Eds.), *Putting Knowledge to Work: New Directions for Knowledge-First Epistemology*. Oxford University Press.

Perner, J. (1991). *Understanding the Representational Mind*. MIT Press.

Perner, J., Huemer, M., & Leahy, B. (2015). Mental files and belief: A cognitive theory of how children represent belief and its intensionality. *Cognition*, *145*, 77–88.

Phillips, J., Buckwalter, W., Cushman, F., Friedman, O., Martin, A., Turri, J., Santos, L., & Knobe, J. (2021). Knowledge before Belief. *Behavioral and Brain Sciences*, *44*(e140), 1–75.

Phillips, J., & Norby, A. (2021). Factive theory of mind. *Mind and Language*, *36*(1), 3–26.

Poulin-Dubois, D., Rakoczy, H., Burnside, K., Crivello, C., Dörrenberg, S., Edwards, K., Krist, H., Kulke, L., Liszkowski, U., Low, J., Perner, J., Powell, L. J., Priewasser, B., Rafetseder, E., & Ruffman, T. (2018). Do infants understand false beliefs? We don't know yet – A commentary on Baillargeon, Buttelmann and Southgate's commentary. *Cognitive Development*, *48*, 302–315.

Poulin-Dubois, D., & Yott, J. (2018). Probing the depth of infants' theory of mind: disunity in performance across paradigms. *Developmental Science*, *21*(4), 1–11. https://doi.org/10.1111/desc.12600

Powell, L. J., Hobbs, K., Bardis, A., Carey, S., & Saxe, R. (2018). Replications of implicit theory of mind tasks with varying representational demands. *Cognitive Development*, *46*(September 2017), 40–50.

Priewasser, B., Rafetseder, E., Gargitter, C., & Perner, J. (2018). Helping as an early indicator of a theory of mind: Mentalism or Teleology? *Cognitive Development*, *46*, 69–78.

Rakoczy, H. (2017). In Defense of a Developmental Dogma: Children Acquire Propositional Attitude Folk Psychology Around Age 4. *Synthese*, *194*(3), 689–707.

Rakoczy, H., & Oktay-Gür, N. (2020). Why Do Young Children Look so Smart and Older Children Look so Dumb on True Belief Control Tasks? An Investigation of Pragmatic Performance Factors. *Journal of Cognition and Development*, *21*(2), 213–239.

Rubio-Fernández, P., & Geurts, B. (2013). How to Pass the False-Belief Task Before Your Fourth Birthday. *Psychological Science*, *24*(1), 27–33.

Sabbagh, M. A., & Paulus, M. (2018). Replication studies of implicit false belief with infants and toddlers. *Cognitive Development*, *46*, 1–3. https://doi.org/10.1016/j.cogdev.2018.07.003

Salter, G., & Breheny, R. (2019). Removing shared information improves 3- and 4-year-olds' performance on a change-of-location explicit false belief task. *Journal of Experimental Child Psychology*, *187*, 104665. https://doi.org/10.1016/j.jecp.2019.104665

Schidelko, L. P., Huemer, M., Schröder, L. M., Lueb, A. S., Perner, J., & Rakoczy, H. (2022). Why Do Children Who Solve False Belief Tasks Begin to Find True Belief Control Tasks Difficult? A

Test of Pragmatic Performance Factors in Theory of Mind Tasks. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.797246

Scholl, B. J., & Leslie, A. M. (1999). Modularity, Development and "Theory of Mind." *Mind & Language*, *14*(1), 131–153.

Schulze, C., & Buttelmann, D. (2022). Infants' differentiate between successful and failed communication in a false-belief context. *Infant Behavior and Development*, *69*. https://orcid.org/0000-0001-9151-4792

Scott, R. M., & Baillargeon, R. (2009). Which penguin is this? Attributing false beliefs about object identity at 18 months. *Child Development*, *80*(4), 1172–1196.

Scott, R. M., Baillargeon, R., Song, H., & Leslie, A. M. (2010). Attributing false beliefs about non-obvious properties at 18 months. *Cognitive Psychology*, *61*(4), 366–395.

Scott, R. M., Roby, E., & Baillargeon, R. (2022). How sophisticated is infants' theory of mind? In O. Houdé & G. Borst (Eds.), *Cambridge handbook of cognitive development* (pp. 242–268). Cambridge University Press. http://labs.psychology.illinois.edu/infantlab/articles/scott_roby_baillargeon_in press.pdf

Scott, R. M., Roby, E., & Setoh, P. (2020). 2.5-Year-Olds Succeed in Identity and Location Elicited-Response False-Belief Tasks With Adequate Response Practice. *Journal of Experimental Child Psychology*, *198*.

Setoh, P., Scott, R. M., & Baillargeon, R. (2016). Two-and-a-half-year-olds succeed at a traditional false-belief task with reduced processing demands. *Proceedings of the National Academy of Sciences of United States of America*, *113*(47), 13360–13365.

Shahaeian, A., Peterson, C. C., Slaughter, V., & Wellman, H. M. (2011). Culture and the Sequence of Steps in Theory of Mind Development. *Developmental Psychology*, *47*(5), 1239–1247.

Shatz, M., Wellman, H. M., & Silber, S. (1983). The acquisition of mental verbs: A systematic investigation of the first reference to mental state. *Cognition*, *14*(3), 301–321. https://doi.org/10.1016/0010-0277(83)90008-2

Siegal, M., & Beattie, K. (1991). Where to look first for children's knowledge of false beliefs. *Cognition*, *38*(1), 1–12.

Sodian, B., Kaltefleiter, L. J., Schuwerk, T., & Kloo, D. (2024). Continuity in false belief understanding from 33 to 52 months of age. *Journal of Experimental Child Psychology*, *247*. https://doi.org/10.1016/j.jecp.2024.106039

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, *18*(7), 587–592.

Southgate, V., & Vernetti, A. (2014). Belief-based action prediction in preverbal infants. *Cognition*, *130*(1), 1–10.

Surian, L., & Leslie, A. M. (1999). Competence and performance in false belief understanding: A comparison of autistic and normal 3-year-old children. *British Journal of Developmental Psychology*, *17*(1), 141–155.

Tauzin, T., & Gergely, G. (2018). Communicative mind-reading in preverbal infants. *Scientific Reports*, *8*(1), 1–9. https://doi.org/10.1038/s41598-018-27804-4

Wellman, H. M. (2014). *Making Minds*. Oxford University Press.

Wellman, H. M., & Bartsch, K. (1988). Young children's reasoning about beliefs. *Cognition*, *30*(3), 239–277.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-Analysis of Theory-of-Mind Development: The Truth about False Belief. *Child Development*, *72*(3), 655–684.

Wellman, H. M., Fang, F., & Peterson, C. C. (2011). Sequential Progressions in a Theory-of-Mind Scale: Longitudinal Perspectives. *Child Development*, *82*(3), 780–792.

Wellman, H. M., & Liu, D. (2004). Scaling of Theory-of-Mind tasks. *Child Development*, *75*(2), 523–541. https://doi.org/10.1111/j.1467-8624.2004.00691.x

Westra, E. (2017). Pragmatic Development and the False Belief Task. *Review of Philosophy and Psychology*, *8*(2), 235–257.

Westra, E., & Carruthers, P. (2017). Pragmatic development explains the Theory-of-Mind Scale. *Cognition*, *158*(1), 165–176.

Westra, E., & Nagel, J. (2021). Mindreading in conversation. *Cognition*, *210*(June 2020), 104618. https://doi.org/10.1016/j.cognition.2021.104618

Williamson, T. (2000). Knowledge and its Limits. In *Knowledge Creation Diffusion Utilization*. Oxford University Press.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 103–128.

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