

COMMENTARY

Should infant psychology rely on the violation-of-expectation method? Not anymore

Markus Paulus 

Department Psychology, Ludwig-Maximilians-Universität München, Munich, Germany

CorrespondenceMarkus Paulus, Department Psychology, Ludwig-Maximilians-Universität München, Leopoldstr. 13, 80802 Munich, Germany.
Email: markus.paulus@lmu.de**Funding information**

James S. McDonnell Foundation, Grant/Award Number: 220020511

Abstract

Infant psychology has heavily drawn on the Violation-of-Expectation (VoE) method. In this piece, I analyse the assumptions that go into the VoE method by comparing it with other looking time measures. Reviewing key arguments in the literature, I argue that this method relies on too many questionable assumptions, that its conceptual foundations are thin, and that its findings are empirically not conclusive. Theoretical inferences based on VoE findings are thus too speculative to allow for solid progress in developmental psychology. Overall, while I appreciate that the VoE method has helped to make infant psychology a prominent field of research, I argue that it is time for infant psychology to move forward and to leave the VoE method behind.

KEYWORDS

cognitive development, habituation, infant psychology, psychological methods, violation-of-expectation

1 | INTRODUCTION

Scientific progress is tightly related to methodological innovation and change of the methods applied in a given field of research. Thereby, agreement on what constitutes a suitable method determines what types of empirical questions can be meaningfully asked and, consequently, how a field of inquiry is constituted (Chalmers, 1999). Thus, methods constrain and guide empirical research. At the same time, methodological views are closely intertwined with (meta-)theoretical frameworks (Witherington, Overton, Lickliter, Marshall, & Narvaez, 2018) as they follow from our basic assumptions about central concepts of a research field. For example, the use of a particle accelerator makes

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2022 The Author. *Infant and Child Development* published by John Wiley & Sons Ltd.

only sense and its results (that are usually difficult to understand for laypersons) are only interpretable if one shares the assumptions of particle physics. Once introduced, it allowed to ask a number of detailed questions that have not been possible to ask before this method has been developed and refined. Likewise, whereas the horoscope was once a respectable method in the science of the universe and human nature, the decay of the astrological framework has undermined its scientific plausibility. Consequently, the questions that can be answered by its application are now meaningless. Although contemporary methods of infant psychology usually lack the precision of particle accelerators and are less speculative than horoscopes, these examples show how deep methods and theoretical progress are related. In a narrower sense, and as explained in the next section, methods can themselves be conceived as being hypotheses that can be evaluated like any other hypothesis. In this essay, I will evaluate the prospects and limitations of one prominent method of infant psychology, the Violation-of-Expectation (VoE) paradigm. Based on this evaluation, I will argue that we should refrain from using this method.

The VoE method relates to a larger set of methods that investigate preverbal children's psychological processes through the assessment of infants' looking times. It builds on pioneering work by Robert Fantz who demonstrated that infants in their first months discriminate between visual patterns by showing clear looking preferences (Fantz, 1958) and show decreased looking to a repeatedly presented stimulus relative to a novel stimulus (Fantz, 1964). The initial methodological approach was used to explore early perceptual development, most notably whether infants can differentiate between two stimuli. Building on the proposal that longer looking time constitutes a surprise reaction to an unexpected (e.g., impossible or rule-violating) event, this approach was then further developed into the VoE method. Based on the proposal that the VoE method allows for an assessment of infants' appreciation of impossible or rule-violating events, its application is supposed to reveal the presence of cognitive competencies in preverbal infants. Its basic idea is that if infants are presented with a, for example, possible and impossible event (as defined by the researcher) and look longer at the impossible event, researchers assume that infants are 'surprised' by the impossible event and then infer that infants have some knowledge of the rules or laws underlying these events. It is assumed that these cognitive competencies are otherwise hidden as infants lack abilities (e.g., motor abilities) to demonstrate them. It is this aspect that will be in the focus of my considerations.

Notably, the VoE paradigm has had a tremendous impact on the field of infant psychology. Traditional accounts had ascribed limited cognitive competencies to young children. For example, Piaget (1952, 1962) described preverbal infants' functioning on a sensorimotor level and proposed it to be tied to the here and now, lacking representational abilities. By using the VoE paradigm, researchers aimed at overcoming a (from their perspective) too pessimistic view of infant competencies. Initially, a large body of VoE research explored physical knowledge in infancy. With the rise of the VoE approach, it has been used to assess a large and ever-growing variety of different competencies. For example, within the research area of naïve physics, it has been proposed that in the first months of life infants have knowledge about physical regularities such as containment, occlusion, and covering (Baillargeon, 2004), and understand aspects of object physics such as cohesion, solidity, and continuity (e.g., Kuhlmeier, Bloom, & Wynn, 2004; Spelke, Breinlinger, Macomber, & Jacobson, 1992; Spelke, Kestenbaum, Simons, & Wein, 1995). In the second half of the first year of life, infants understand the principle of inertia (Spelke, Katz, Purcell, Ehrlich, & Breinlinger, 1994). Concerning intuitive psychology, it has been proposed that infants expect agents to behave rationally (Gergely, Nadasdy, Csibra, & Biro, 1995; cf. Gergely & Csibra, 2003), understand human as intentional actors (Phillips & Wellman, 2005) and 'intentions as existing independently of particular concrete actions and as residing within the individual' (Woodward, 2009, p. 55), understand the complementary nature of actions if people work toward a collaborative goal (Henderson & Woodward, 2011), and understand others' false beliefs (Kovacs, Teglas, & Endress, 2010). More recent research focused on infants' understanding of the social world. It has been proposed that infants have some understanding of the conventionality of human language use (Novack, Henderson, & Woodward, 2014), appreciate social dominance (Thomsen, Frankenhuis, Ingold-Smith, & Carey, 2011), the stability of social dominance (Mascaro & Csibra, 2012), and use group size to infer dominance relationships between social groups (Pun, Birch, & Baron, 2016), have an understanding of fairness principles (e.g., Ziv & Sommerville, 2017), attribute moral traits to agents (Surian, Ueno, Itakura, & Meristo, 2018), understand ingroup dynamics (Ting, He, &

Baillargeon, 2019) and expect ingroup issues to override fairness concerns (Bian, Sloane, & Baillargeon, 2018). Further research programs demonstrated numerical competencies (Wynn, 1995) and logical reasoning abilities (Cesana-Arlotti, Kovacs, & Teglas, 2020) in infants. The wide popularity of the VoE approach (and related measures) might be due to the fact that it is a rather inexpensive method that is easy to implement and that allows for an assessment of any possible knowledge domain. Overall, studies using the VoE (and related) paradigms have revolutionized the field's picture of infant development. It assumes cognitive capacities or knowledge that is otherwise not detectable as, so the story goes, infants might not be able to demonstrate their knowledge (e.g., due to a lack of motor control, or linguistic competencies). That is, they are supposed to possess 'impressive early competence' (Wynn, 1995, p. 172) and 'sophisticated knowledge' (Willatts, 1997, p. 112) that has been hidden from us as it has so far only been observed by the (late) preschool period, and that can be revealed by the VoE approach.

2 | PUTTING THE VOE PARADIGM IN PERSPECTIVE

Certainly, the strength and conclusiveness of a method is not a simple all-or-nothing issue but calls for a more nuanced and fine-grained evaluation. To do so, we can gain valuable insights by comparing different approaches and place the method into a continuum of related approaches. This will allow us to gain a better appreciation of the conclusiveness of this method. In the following, I will therefore compare the VoE method to other methods relying on measures of young children's looking behaviour. Hereby, I follow the classification offered by Houston-Price and Nakai's (2004) comprehensive review and will focus on the Preferential looking method, the Habituation-dishabituation method, the Familiarization-novelty preference method, and the VoE method. I will evaluate each method with respect to the additional assumptions that researchers make when using it, and thus the conclusiveness of the inferences drawn from it.

The classical 'Preferential looking method' describes an approach in which infants are presented simultaneously with two stimuli and the amount of looking at each stimulus is compared. This method allows thus to assess whether infants perceive a difference between the two stimuli and whether (and to which extent) they prefer one stimulus over the other. For example, Bushnell (2001) presented a few hours old infants with their mother's and a stranger's face, and assessed their looking preference. The study showed that infants' preference for the mother increased with the amount of exposure to their mother's face. Whereas infants with very little exposure showed no systematic preference, this preference was established after a few hours of exposure to their mother. From a methodological point of view, I would highlight two points. First, one assumption is that there is a clear linear relation between the amount of looking and the amount of attention—notably, as all the other methods draw on looking time measures, this seems to be common to all of them. Second, the 'Preferential looking method' also has limitations as null results are inconclusive: infants might perceive a difference, but due to a lack of (systematic) preference may not show a difference in looking time, or infants might not perceive a difference at all. Yet, overall I regard this paradigm as least controversial as the inferences drawn from it (e.g., infants perceive a difference, they look longer at one stimulus over the other) are quite closely related to the measure itself. That is, it does not require a lot of additional assumptions to connect the measure with the theoretical construct as, for example, the concept of visual preference is tightly related to a greater dedication of visual attention to one entity over the other.

The 'Habituation-dishabituation method' goes one step further. Here, infants are first repeatedly presented with a stimulus. This usually leads to a decline of looking time over the course of successive stimulus presentations. This phenomenon is known as habituation and is seen as one of the most basic forms of learning (Colombo & Mitchell, 2009). After repeated exposure, a different stimulus is presented. An increase of looking time to the novel stimulus ('dishabituation') is supposed to indicate that infants perceived a difference between the stimuli. It is particularly useful when no a priori preference for one of the stimuli exists and a 'Preferential looking method' would thus not be helpful (Houston-Price & Nakai, 2004). In fact, the habituation phase induces satiation with the old stimulus

and thus aims for an experimentally induced visual preference for the novel stimulus—if a difference is perceived. By systematically manipulating the features along which the two stimuli differ, it might indicate which aspects of the old stimulus have been encoded. For example, Oakes and Cohen (1990) did not find any a priori preferences in 6- and 10-month-old infants' looking to causal and non-causal events. Yet, after being habituated to either of the two event classes, the 10-month-old infants dishabituated to the respective other event class indicating that they perceive causal events differently than non-causal events. From a methodological point of view, it needs to be noted that differences in looking time to the old and novel stimulus require some recognition memory (Houston-Price & Nakai, 2004). Inferences drawn are from findings following this method thus rely on additional assumptions about learning and memory processes. Furthermore, it can be debated how long infants have to be habituated to the original stimulus—one rigorous way is to use a so-called infant-controlled procedure in which habituation is determined individually once an infants' looking time has decreased according to a beforehand set criterion (Cohen, 2004)—and what it means if some infants do not reach the habituation criterion. In addition, when comparing types of stimuli (e. g., causal vs. non-causal events) scholars aim at inferences that might go beyond the perceptual level. Therefore, it is more complex in its assumptions than the 'Preferential looking method'.

A further step is done by the 'Familiarization-novelty preference method' that combines elements of the two previously described approaches (Houston-Price & Nakai, 2004). Similar to the 'Habituation-dishabituation method', infants are first familiarized or habituated to one stimulus (or a pair of the same stimulus). Thereafter, following the 'Preferential looking method', a (novel) pair of stimuli are presented in a test phase; the original stimulus and the novel stimulus. Infants' preferential looking to the novel stimulus is proposed to indicate their ability to differentiate between the stimuli. One could argue that the simultaneous presentation of the original and the novel stimulus reduces memory demands (compared to Habituation-dishabituation method in which the original stimulus is not presented) while still allowing to rely on the advantages of the habituation procedure. Yet, some studies (e.g., Jusczyk & Aslin, 1995; Roder, Bushnell, & Sasseville, 2000) did not reveal clear novelty preferences, but also preferences for the original stimulus (familiarity preferences), which seems to be surprising at first sight. This has been explained by the idea that an insufficient encoding of the original stimulus in the first phase leads to an inclination to complete encoding in the test phase—and hence a preference for the original stimulus. Once encoding is completed, infants start to prefer to look to the novel stimulus (Hunter & Ames, 1988). That is, specific theoretical models have been developed that try to explain the time course of infants' preference for the familiar or novel stimulus in relation to the familiarization time. Yet, given considerable individual variability, it is difficult to estimate the presence of familiarity and novelty effects on a group level (Bogartz & Shinskey, 1998; see also Bogartz, Shinskey, & Speaker, 1997). Furthermore, the empirical support for the theoretical models is not as clear as one would wish (e.g., Bergmann & Cristia, 2016). Slater (2004) concedes that 'this attentional shift is not always found, and that in many instances we need converging evidence in order to clarify experimental findings' (p. 353). That is, both directions of effect are possible (familiarity preference, novelty preference), their appearance may sometimes not be easy to predict, and their theoretical meaning is subject to debate. If one indeed needs 'to find converging evidence for what may turn out to be the 'correct' interpretation of the findings' (Slater, 2004, p. 355), it seems fair to state that this method relies on a number of additional assumptions. Interpretations based on this paradigm seem to be less conclusive and more speculative than those of other paradigms.

Finally, let us turn to the VoE paradigm. In one prominent version (e.g., Baillargeon, Spelke, & Wasserman, 1985), infants are first habituated to a stimulus (often an entire event or scenario) until a preset habituation criterion is reached. The stimulus usually is supposed to be coherent with what we know about the world. In the test phase, infants are presented with pairs of test trials. Importantly, instead of comparing infants' preference for a familiar versus a novel stimulus, they are presented with two novel stimuli that both differ from the original stimulus (leading to considerations to which extent the habituation or familiarization stimuli are actually necessary for VoE tasks; for example, Wang, Baillargeon, & Brueckner, 2004). One of the test events is supposed to be perceptually more novel but not 'impossible' and thus not violating any expectation. The other event is supposed to be perceptually more similar to the familiarization event (that is, sharing crucial perceptual features), but 'impossible' or highly unlikely with respect to our

standards (e.g., a physical law, a social rule, or an otherwise impossible or highly unlikely event). If infants have the respective knowledge, this event should be surprising and infants should thus look longer at this event than at the other event. In other words, if infants look longer to this event, it is concluded that this stimulus attracted infants' attention as they recognized its impossibility or otherwise rule-violating nature. Hence, it is indicative for infants' cognitive capacities or knowledge, respectively.

One common concern is that the characterization of the test trials relies on assumptions on the novelty or familiarity of the different events with respect to the familiarization stimulus. To the extent that the events contain arrays of different objects and movements, it is difficult to determine a priori on which aspects of the events infants actually focus—and thus, what actually constitutes familiarity and novelty in these contexts. This ties in with several studies that (as will be discussed in more detail in the next section) suggested that sometimes subtle perceptual changes may lead to very different pattern of results (e.g., Schöner & Thelen, 2006)—thus implying an interpretation of VoE paradigms on a perceptual level (Haith, 1998). It should be noted, though, that some studies include control conditions with the aim to exclude simple perceptual interpretations (Munakata, 2000). Yet, given that the scenarios often contain an abundance of features that could be the focus of infants' processing and that cannot all be controlled for, theoretical inferences based on VoE findings do not seem very conclusive.

This becomes even more complicated in studies in which the events consist of complex scenarios that involve a lot of information—that is, perceptually complex scenarios in which human or animated agents pursue different actions whose meaning may depend on cultural practices or are only rudimentary sketched (e.g., Onishi & Baillargeon, 2005; Powell & Spelke, 2013; Tatone & Csibra, 2015). Researchers need to make a considerable number of assumptions on how infants process all of this information, that they understand all of the (from the researcher's point of view) relevant aspects, and that they indeed focus on the key variables of interest. From a theoretical point of view, we need to consider to what extent these assumptions are justified and the original interpretation is thus rational (cf. Paulus, 2012). For example, in Onishi and Baillargeon's (2005) study on infant false belief understanding, 15-month-old infants first observed an actor with a visor putting an object (a toy watermelon slice) in one of two boxes and then reaching two times into this box (familiarization). In the two test trials, they showed how the protagonist reached in each of the two boxes, respectively. Between familiarization and test trials, they presented participants with different types of events that aimed at inducing the protagonist's true or false belief. The events consisted of, for example, the boxes moving by itself, the object moving by itself from one box into the other box while the protagonist was either present or absent. This paradigm requires participants to understand several issues, for example, what the protagonist with the visor can and cannot see. At the same time, one needs to assume that participants do not focus on other (quite surprising) characteristics, for example that the boxes move by themselves forth and back, or that toy watermelon slices can wander around by themselves. Moreover, it is unclear how to interpret that the protagonist reaches two times into the box in which the toy was previously hidden without actually taking it out (familiarization). It does not clearly indicate that she searches for the toy. Yet, the authors assume that infants interpret this behaviour as searching for the toy and that infants also conceive of the test trials in the same way—a necessary prerequisite to interpret the test trials as true or false belief trials. Overall, the study involves a number of methodological decisions that are difficult to find strong justifications for (Poulin-Dubois et al., 2018). Thus, this example shows the extent to which VoE paradigms rest on a considerable number of additional assumptions that might in themselves be questionable. Taken together, inferences drawn from this paradigm are more speculative and less conclusive than those from the other paradigms.

Overall, the closer analysis of the different looking time paradigms revealed that the approaches differ considerably in the number of additional assumptions that need to be made. As a consequence, the conclusiveness of the findings and the strength of the inferences that can be drawn from the empirical results vary considerably (see Figure 1) with the VoE-paradigm being the most speculative and least conclusive of the here considered approaches.

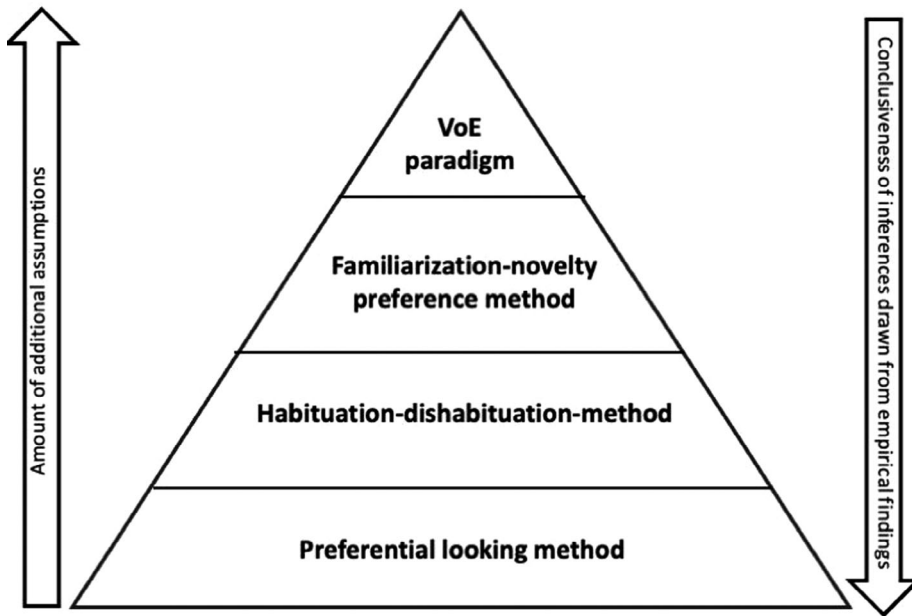


FIGURE 1 Relation between different looking-time paradigms, the amount of assumptions they come with, and the conclusiveness of the inferences that can be drawn from them

3 | METHODS ARE HYPOTHESES

Our thoughts about the usefulness and applicability of a method are sentences that can be evaluated—and even falsified by falsification of their corollaries (Popper, 2005). In short, we can say that methods can be conceived as being hypotheses. Indeed, following a deductive-nomological model of scientific explanation and hypothesis generation, the prediction of concrete study results usually relies on a (large and often unspecified) number of assumptions (the Duhem-Quine thesis; Chalmers, 1999). To the extent that the assumptions are potentially falsifiable, we can conceive of them as auxiliary hypotheses. One of the key auxiliary hypotheses is the usefulness, applicability, and conclusiveness of the respective method. Usually, we assume that our ‘method works’—probably often based on the consideration that it has been ‘successfully’ used in other studies. Nonetheless, it is good to pause for a moment and to reflect on the strengths of our auxiliary hypotheses. This is an important issue as, for example, whenever there is an undesired result (usually: a null finding), one needs to decide whether the original hypothesis is unlikely (that is, the negative result falsifies the original theory) or whether the auxiliary hypothesis is unlikely (that is, the method is not working well). Thus, we have to ask ourselves how strongly we are convinced by the VoE method.

Notably, there is an ongoing dispute on the conclusiveness and the validity of the VoE approach (e.g., Aslin, 2007; Haith, 1998; Tafreshi, Thompson, & Racine, 2014). There is disagreement on what is actually measured when relying on VoE (Tafreshi et al., 2014) and on whether we really tap into cognitive processes (e.g., Bogartz et al., 1997; Cohen, 2004; Schilling, 2000; Schöner & Thelen, 2006). In the following paragraphs, I will first outline conceptual problems associated with the VoE paradigm before describing empirical considerations and turning to questions of validity. These thoughts tie in with the considerations put forward in the previous section.

From a conceptual point of view, Tafreshi et al. (2014) highlighted that (longer) looking time in natural and laboratory contexts can be an indicator of a variety of psychological states such as fear, preference, anticipation, surprise, annoyance, or scepticism. Whereas in natural contexts the meaning of looking behaviour can be clarified by asking the respective person and/or by observing the wider natural context in which the behaviour is embedded, and can

then be meaningfully interpreted, this is not the case in laboratory contexts. We are thus faced with conceptual problems when ascribing a particular meaning to infant-looking behaviour, that is, 'conceptual problems emerge when one insists that a capacity for change detection requires ascriptions of surprise or preference, let alone understanding of belief' (Tafreshi et al., 2014, p. 230). Notably, this is a conceptual question in which the meaning of a particular behaviour (here: looking time) is under dispute. This conceptual question needs to be clarified before the measure can be meaningfully employed in an empirical study.

The conceptual gap between method and inference becomes clearer when considering the inferential chain. Following Tafreshi et al.'s (2014) analyses, I will highlight two steps in the chain, both of which are highly questionable and very thin. The main point here is that the VoE method is not used to explore the factors related to infant-looking behaviour (which might be less of a problem), but to make a number of additional inferences. First, in the VoE paradigm, longer looking is equated with surprise. Yet, there is no unequivocal relation between longer looking and surprise. We would be very hesitant to ascribe surprise to any instance in which a person looks long or stares at something. Individuals' looking behaviour might be indicative of a variety of other processes such as fear, hate, desire, or disgust. Likewise, we might ascribe the state of surprise to someone even when this individual does not show enhanced looking. Thus, the link between longer looking and surprise is rather weak. Second, another step in the inferential chain involves inferences from assumed surprise to underlying cognitive insights. Individuals can have cognitive insights or engage in cognitive activities (e.g., making moral judgements; ascribe false beliefs; recognize social hierarchies) without being surprised. Likewise, being surprised is not an unequivocal indicator of particular cognitive capacities or insights, let alone the consideration of fairness principles, the ascription of particular mental states to others or the capacity for logical inferences. Thus, the link between (assumed) surprise and cognitive capacities is rather weak. Taken together, there is a wide gap between method and inference with the inferential chain standing on unsound or at least on shaky grounds.

From an empirical point of view, it can be asked whether the results of VoE studies can be explained by relying on simple perceptual processes without the need of introducing advanced cognitive capacities. Haith (1998) had argued that in looking time paradigms all possible perceptual factors need to be controlled for before introducing additional cognitive concepts. One prominent example of this empirical criticism was provided by Schöner and Thelen (2006). They presented a dynamic field model on infants' habituation and dishabituation reactions, and used this approach to model results of previous studies that claimed object knowledge in 3.5- to 5-month-old children (Baillargeon, 1987; Baillargeon et al., 1985). Interestingly, their model could account for these findings by simply modelling activation and inhibitory processes – without invoking any kind of conceptual understanding. One implication of the model is that "small changes in the stimuli, in the timing of presentations, in the metric differences of the test events, in the number of trials and habituation criteria, and in individual infants can radically change the outcome measure" (Schöner & Thelen, 2006, p. 274). And indeed, it was able to predict result patterns of previous studies such as unexpected order effects that had been puzzling for researchers. That is, a variety of findings could be explained without the need to ascribe conceptual understanding to infants (Sirois & Jackson, 2007; see also Uithol & Paulus, 2014). Thus, we need to consider whether we are too quickly committing ourselves to conceptual interpretations of phenomena that might be more readily explained on an associative level. This seems to be a general risk for child psychology. Wilhelm Wundt famously stated: 'Child-psychology often suffers from the same mistake that is made in animal psychology: namely, that the observations aren't interpreted objectively, but are filled out with subjective reflections. Thus, the earliest ideational combinations, which are in reality purely associative, are regarded as acts of logical reflection [...]'. (Wundt, 1897/1997). Taken together, it seems that the key promise that spurred the application of the VoE approach (that is, revealing hidden cognitive competencies), cannot be kept without doubts, and even on empirical grounds the inferences drawn from the VoE approach are not very conclusive.

Similarly, while researchers put a lot of effort and interpretation in the construction of stimuli that, in order to test cognitive understanding (and not merely perceptual differentiation), need to be conceptually rich in content, it is unclear how exactly children understand the scenarios. Partly, the stimuli consist of several sequences of actions. They seem to be intuitive for those who work in a particular research area but are sometimes even challenging for

adults without a specific background in psychological science. On a more general level, the problem is famously known as ‘psychologist’s fallacy’ (James, 1981/1890). One example comes from a study that presented adults with three video scenarios that had been used in VoE tasks in order to examine false belief understanding in infancy (Low & Edwards, 2018). Notably, only one of the three scenarios was actually deemed plausible by the adults—leading thus to questions how to interpret the respective findings in child populations.

On a side note, it should be considered that, in contrast to the initial development of looking time measures, the main point of most VoE studies is not to examine whether children are able to perceptually differentiate between two stimuli, but to draw quite specific inferences about their cognitive capacities. Yet, in contrast to the specificity of the conclusions, the VoE is a rather unspecific measure. It merely shows that infants make a difference between events. There is thus a considerable gap between the specificity of conclusions that one would like to draw and the rather unspecific nature of the measure. One could speculate that the unspecific nature of the measure has all too often limited developmental research to merely inquiry whether infants have particular competencies and whether it can be found in even younger children if the task is simplified—instead of striving for explanatory models (see also Hommel, 2020) and instead of explaining human development. Notably, if one is interested in studying infant-looking behaviour, recent methodological developments allow for a much more differentiated assessment of infant-looking behaviour. For example, eye-tracking allows to assess where infants look, how they scan the visual information (e.g., whether infants actually look at the—from a theoretical point of view—most relevant aspects of the scene), and how their gaze at specific parts of the scene changes across time (Gredebäck, Johnson, & von Hofsten, 2010). Notably, I do not want to argue that, for example, eye-tracking is devoid of additional assumptions—indeed, it involves assumptions on the technical side as well as on the suitability of the stimuli. Nonetheless, its greater spatial and temporal precision, and the lack of reliance on putative cognitive processes (i.e., longer looking = surprise = violation of a cognitive expectation)—has the potential to make it a less speculative measure. One keeps wondering why we should stick to a rather unspecific method if more specific and powerful assessment tools are available.

This directly speaks to the third point, that is, how can we know about the validity of the VoE approach. If we consider the validity of other measures in which competencies are assessed, we develop the idea of assessing these competencies by observing them in real life. That is, when deriving a measure of intelligence, it is because we observe intelligent behaviours in real life and we deem them to be a valuable topic for scientific inquiry. Likewise, measures for pro- and antisocial behaviours are developed because we are faced with pro- and antisocial behaviours and we wonder about the factors promoting or preventing them, about the right way to assess them, and their consequences. This is quite different for findings or conclusions derived from VoE approaches. When interacting with infants, there is no evidence that they reason about true or false beliefs. What does it then mean to state that infants understand others’ false beliefs (Kovacs et al., 2010)? Similarly, infants do not form stable groups or engage in group activities, they do not establish hierarchical relations in groups. What does it then mean to state that infants understand social dominance as a stable relation (Mascaro & Csibra, 2012)? This list could be extended by a large number of ascribed and by means of the VoE approach apparently demonstrated competencies that are otherwise not observable in infants’ behaviour. How can infants have all of these competencies that are seemingly not apparent in their actual behaviour and ‘real life’? One could speculate that the rather easy and inexpensive implementation of the VoE approach and its independence of any content-related considerations virtually invites to search for any conceivable cognitive competence in infants without the need for validating it against real behaviour.

It should be mentioned that there is also direct evidence for a dissociation between the competencies ascribed to infants and the lack of these competencies in older children. Studies with older children that used more conclusive measures (e.g., active search behaviour; verbal responses) to study the competencies that had been ascribed to preverbal infants, sometimes failed to show such understanding in children who are partly even several years older (e.g., Gönül & Paulus, 2021; Hinten, Labuschagne, Boden, & Scarf, 2018; Hood, Carey, & Prasada, 2000; Krist, Karl, & Krüger, 2016). Moreover, other studies used different methods with preverbal infants (e.g., eye-tracking) and partly found different results (e.g., Ganglmayer, Attig, Daum, & Paulus, 2020). Of course, it is possible that the VoE method is a much more sensitive and conclusive measure for children’s real competencies. Indeed, it has been suggested that

infants' true competencies are masked by other 'more demanding' measures as they rely on additional abilities or that these competencies are hidden behind further psychological processes that emerge later in development (e.g., Krist et al., 2016). Yet, how could one ever prove this if any failure to find converging evidence can be dismissed by suggesting that all other measures are just too demanding. Thus, given all the above-mentioned considerations with respect to the conclusiveness of the VoE method, should we really bet on that?

Relatedly, the question arises which anthropological model is implicated in the VoE method and how valid is this model? Notably, this view of infants seems to be quite similar to patients suffering from a locked-in syndrome. This syndrome describes a disorder in which persons are conscious (as opposed to having lost consciousness), still dispose of their knowledge but cannot demonstrate it or their will as they lack control over the effectors including their ability to use language. The major way to express themselves is by moving their eyes. Interestingly, there is some analogy to the picture inherent to the VoE approach: Infants cannot show their rich knowledge or competencies due to a lack of motor control and language abilities. Certainly, it is a different situation as locked-in patients have already acquired propositional language whereas infants did not. Nonetheless, there is some analogy in the characterization of infants and patients suffering from a locked-in syndrome (I thus call it the *locked-in syndrome view of infants*). One could wonder whether this is a valid picture of human infancy.

In sum, even decades after it has been introduced and major, impactful studies have been published (e.g., Baillargeon et al., 1985; Spelke et al., 1992, 1995) and keep being published (e.g., Liu, Brooks, & Spelke, 2019; Mascaro & Csibra, 2012), the VoE approach is controversial and, most of all, its conceptual basis remains shaky (Tafreshi et al., 2014). Given the above-mentioned concerns, would it then not be better to refrain from using this method? To the extent that the auxiliary hypothesis is doubtful, we might risk putting strong theoretical claims on sandy land instead of on solid rock.

4 | CONCLUSION

The use of habituation-based measures has been gone beyond what they were initially intended to study (Fantz, 1964) and what seemed to allow for reasonable inferences (Cohen, 2004). This essay suggests that the VoE approach suffers from too many limitations and questionable assumption so that it seems reasonable to terminate its use. One might argue that an alternative way would be to refine and improve the VoE methodology in order to deal with the concerns. Yet, the fundamental nature of the problems as discussed above do not allow for a simple technological adjustment, they concern the very nature of the approach.

Notably, this piece was written as an essay. It aimed to be concise and pointed. Let me be conciliatory at the end. VoE had its time. It contributed to the visibility of developmental research within psychology and initiated central debates in the field. It seems fair to say that infant psychology would not have become such a rich field without the debates initiated by VoE studies. But now it is time to move forward and to build our theories on more conclusive grounds. As we all know, development involves both gains and losses (Baltes, 1987). In other words, moving forward is not possible without leaving things behind. Let's move on!

ACKNOWLEDGEMENTS

For valuable feedback on a previous version, I am thankful to Christina Bergmann, Natalie Christner, Nina-Alisa Hinz, Maria Mammen, and Antonia Misch.

Preparation of this manuscript was supported by a James S. McDonnell Foundation 21st Century Science Initiative in Understanding Human Cognition–Scholar Award (No. 220020511).

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Markus Paulus  <https://orcid.org/0000-0003-0446-4956>

REFERENCES

- Aslin, R. N. (2007). What's in a look? *Developmental Science*, *10*, 48–53.
- Baillargeon, R. (1987). Object permanence in 3½- and 4½-month-old infants. *Developmental Psychology*, *23*, 655–664.
- Baillargeon, R. (2004). Infants' physical world. *Current Directions in Psychological Science*, *13*, 89–94.
- Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition*, *20*, 191–208.
- Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental Psychology*, *23*, 611–626.
- Bergmann, C., & Cristia, A. (2016). Development of infants' segmentation of words from native speech: A meta-analytic approach. *Developmental Science*, *16*, 901–917.
- Bian, L., Sloane, S., & Baillargeon, R. (2018). Infants expect ingroup support to override fairness when resources are limited. *Proceedings of the National Academy of Sciences of the United States of America*, *115*, 2705–2710.
- Bogartz, R. S., Shinsky, J. L., & Speaker, C. J. (1997). Interpreting infant looking: The event set × event set design. *Developmental Psychology*, *33*, 408–422.
- Bogartz, R., & Shinsky, L. (1998). On perception of a partially occluded object in 6-month-olds. *Cognitive Development*, *13*, 141–163.
- Bushnell, I. W. R. (2001). Mother's face recognition in newborn infants: Learning and memory. *Infant and Child Development*, *10*, 67–74.
- Cesana-Arlotti, N., Kovacs, A. M., & Teglas, E. (2020). Infants recruit logic to learn about the social world. *Nature Communications*, *11*, 5999.
- Chalmers, A. F. (1999). *What is this thing called science?*. St. Lucia: University of Queensland Press.
- Cohen, L. B. (2004). Uses and misuses of habituation and related preference paradigms. *Infant and Child Development*, *13*, 349–352.
- Colombo, J., & Mitchell, D. W. (2009). Infant visual habituation. *Neurobiology of Learning and Memory*, *92*, 225–234.
- Fantz, R. L. (1958). Pattern vision in young infants. *The Psychological Record*, *8*, 43–47.
- Fantz, R. L. (1964). Visual experience in infants: Decreased attention to familiar patterns relative to novel ones. *Science*, *146*, 668–670.
- Ganglmayer, K., Attig, M., Daum, M. M., & Paulus, M. (2020). Infants' perception of goal-directed actions: A multi-lab replication reveals that infants anticipate paths and not goals. *Infant Behavior and Development*, *57*, 101340.
- Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naive theory of rational action. *Trends in Cognitive Sciences*, *7*, 287–292.
- Gergely, G., Nadasdy, Z., Csibra, G., & Biro, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*, 165–193.
- Gönül, G., & Paulus, M. (2021). Children's reasoning about the efficiency of others' actions: The development of rational action prediction. *Journal of Experimental Child Psychology*, *204*, 105035.
- Gredebäck, G., Johnson, S., & von Hofsten, C. (2010). Eye tracking in infancy research. *Developmental Neuropsychology*, *35*, 1–19.
- Haith, M. M. (1998). Who put the cog in infant cognition: Is rich interpretation too costly? *Infant Behavior and Development*, *21*, 167–179.
- Henderson, A. M. E., & Woodward, A. L. (2011). “Let's work together”: What do infants understand about collaborative goals? *Cognition*, *121*, 12–21.
- Hinten, A. E., Labuschagne, L. G., Boden, H., & Scarf, D. (2018). Preschool children and young adults' preferences and expectations for helpers and hinderers. *Infant and Child Development*, *27*, e2093.
- Hommel, B. (2020). Pseudo-mechanistic explanations in psychology and cognitive neuroscience. *Topics in Cognitive Science*, *12*, 1294–1305.
- Hood, B., Carey, S., & Prasada, S. (2000). Predicting the outcomes of physical events: Two-year-olds fail to reveal knowledge of solidity and support. *Child Development*, *71*, 1540–1554.
- Houston-Price, C., & Nakai, S. (2004). Distinguishing novelty and familiarity effects in infant preference procedures. *Infant and Child Development*, *13*, 341–348.
- Hunter, M., & Ames, E. (1988). A multifactor model of infant preferences for novel and familiar stimuli. In C. Rovee-Collier & L. Lipsitt (Eds.), (pp. 69–95). Stamford: Ablex.
- James, W. (1981). *The principles of psychology*. Cambridge: Harvard University Press (Original published 1890).
- Jusczyk, P., & Aslin, R. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, *29*, 1–23.

- Kuhlmeier, V. A., Bloom, P., & Wynn, K. (2004). Do 5-month-old infants see humans as material objects?. *Cognition*, *94*, 95–103.
- Kovacs, A. M., Teglas, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science*, *330*, 1830–1834.
- Krist, H., Karl, K., & Krüger, M. (2016). Contrasting preschoolers' verbal reasoning in an object-individuation task with infants' preverbal feats. *Cognition*, *157*, 205–218.
- Liu, S., Brooks, N. B., & Spelke, E. S. (2019). Origins of the concepts cause, cost, and goal in prereaching infants. *PNAS*, *116*, 17747–17752.
- Low, J., & Edwards, K. (2018). The curious case of adults' interpretations of violation-of-expectation false belief scenarios. *Cognitive Development*, *46*, 86–96.
- Mascaro, O., & Csibra, G. (2012). Representation of stable social dominance relations by human infants. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 6862–6867.
- Munakata, Y. (2000). Challenges to the violation-of-expectation paradigm: Throwing the conceptual baby out with the perceptual processing bathwater? *Infancy*, *1*, 471–477.
- Novack, M. A., Henderson, A. M. E., & Woodward, A. L. (2014). Twelve-month-old infants generalize novel signed labels, but not preferences across individuals. *Journal of Cognition and Development*, *15*, 539–550.
- Oakes, L. M., & Cohen, L. B. (1990). Infant perception of a causal event. *Cognitive Development*, *5*, 193–207.
- Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*, 255–258.
- Paulus, M. (2012). Is it rational to assume that infants imitate rationally? A theoretical analysis and critique. *Human Development*, *55*, 107–121.
- Phillips, A. T., & Wellman, H. M. (2005). Infants' understanding of object-directed action. *Cognition*, *98*, 137–155.
- Piaget, J. (1952). *The origins of intelligence in children*. New York: International University Press.
- Piaget, J. (1962). *Play, dreams and imitation in childhood*. New York: Norton.
- Popper, K. R. (2005). In H. Keuth (Ed.), *Logik der Forschung* (11th ed.). Tübingen: Mohr Siebeck.
- Poulin-Dubois, D., Rakoczy, H., Burnside, K., Crivello, C., Dörrenberg, S., Edwards, K., ... 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.
- Powell, L. J., & Spelke, E. S. (2013). Preverbal infants expect members of social groups to act alike. *PNAS*, *110*, E3965–E3972.
- Pun, A., Birch, S. A. J., & Baron, A. S. (2016). Infants use relative numerical group size to infer social dominance. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, 2376–2381.
- Roder, B., Bushnell, E., & Sasseville, A.-M. (2000). Infants' preferences for familiarity and novelty during the course of visual processing. *Infancy*, *1*, 491–508.
- Schilling, T. (2000). Infants' looking at possible and impossible screen rotations: The role of familiarization. *Infancy*, *1*, 389–402.
- Schöner, G., & Thelen, E. (2006). Using dynamic field theory to rethink infant habituation. *Psychological Review*, *113*, 273–299.
- Sirois, S., & Jackson, I. (2007). Social cognition in infancy: A critical review of research on higher-order abilities. *European Journal of Developmental Psychology*, *4*, 46–64.
- Slater, A. (2004). Novelty, familiarity, and infant reasoning. *Infant and Child Development*, *13*, 353–355.
- Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, *99*, 605–632.
- Spelke, E. S., Katz, G., Purcell, S. E., Ehrlich, S. M., & Breinlinger, K. (1994). Early knowledge of object motion: Continuity and inertia. *Cognition*, *51*, 131–176.
- Spelke, E. S., Kestenbaum, R., Simons, D., & Wein, D. (1995). Spatiotemporal continuity, smoothness of motion and object identity in infancy. *The British Journal of Developmental Psychology*, *13*, 113–142.
- Surian, L., Ueno, M., Itakura, S., & Meristo, M. (2018). Do infants attribute moral traits? Fourteen-month-olds' expectations of fairness are affected by agents' antisocial actions. *Frontiers in Psychology*, *9*, 1649.
- Tafreshi, D., Thompson, J. J., & Racine, T. P. (2014). An analysis of the conceptual foundations of the infant preferential looking paradigm. *Human Development*, *57*, 222–240.
- Tatone, D., & Csibra, G. (2015). Giving and taking: Representational building blocks of active resource-transfer events in human infants. *Cognition*, *137*, 47–62.
- Thomsen, L., Frankenhuys, W. E., Ingold-Smith, M. C., & Carey, S. (2011). Big and mighty: Preverbal infants mentally represent social dominance. *Science*, *331*, 477–480.
- Ting, F., He, Z., & Baillargeon, R. (2019). Toddlers and infants expect individuals to refrain from helping an ingroup victim's aggressor. *Proceedings of the National Academy of Sciences of the United States of America*, *116*, 6025–6034.
- Uithol, S., & Paulus, M. (2014). What do infants understand of others' action? A theoretical account of early social cognition. *Psychological Research*, *78*, 609–622.

- Wang, S., Baillargeon, R., & Brueckner, L. (2004). Young infants' reasoning about hidden objects: Evidence from violation-of-expectation tasks with test trials only. *Cognition*, *93*, 167–198.
- Willatts, P. (1997). Beyond the "couch potato" infant: How infants use their knowledge to regulate action, solve problems, and achieve goals. In G. Bremner, A. Slater, & G. Butterworth (Eds.), *Infant development: Recent advances* (pp. 109–135). London: Taylor & Francis.
- Witherington, D. C., Overton, W. F., Lickliter, R., Marshall, P. J., & Narvaez, D. (2018). Metatheory and the primacy of conceptual analysis in developmental science. *Human Development*, *61*, 181–198.
- Woodward, A. L. (2009). Infants' grasp of others' intentions. *Current Directions in Psychological Science*, *18*(1), 53–57.
- Wundt, W. (1997). *Outline of psychology* (transl. C.H. Judd, 1897). Retrieved April 30, 2021, from York University (Canada), Classics in the History of Psychology Web site: Retrieved from <http://psychclassics.yorku.ca/Wundt/Outlines/sec20.htm> (Original published 1896)
- Wynn, K. (1995). Infants possess a system of numerical knowledge. *Current Directions in Psychological Science*, *4*, 172–177.
- Ziv, T., & Sommerville, J. A. (2017). Developmental differences in infants' fairness expectations from 6 to 15 months of age. *Child Development*, *88*, 1930–1951.

How to cite this article: Paulus, M. (2022). Should infant psychology rely on the violation-of-expectation method? Not anymore. *Infant and Child Development*, e2306. <https://doi.org/10.1002/icd.2306>