Developmental Psychology
2012, Vol. 48, No. 4, 1124–1132

0012-1649/11/$12.00 DOI: 10.1037/a0025915

Three-Year-Olds Understand Appearance and Reality—Just Not About the Same Object at the Same Time

Henrike Moll and Michael Tomasello
Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

Young children struggle in the classic tests of appearance versus reality. In the current Study 1, 3-year-olds had to determine which of 2 objects (a deceptive or a nondeceptive one) an adult requested when asking for the "real X" versus "the one that looks like X." In Study 2, children of the same age had to indicate what a single deceptive object (e.g., a chocolate-eraser) looked like and what it really was by selecting one of two items that represented this object's appearance (a chocolate bar) or identity (a regular eraser). Children were mainly successful in Study 1 but not in Study 2. The findings are discussed with a focus on young children's difficulty with "confronting" perspectives, which may be involved in their struggles with a number of classic theory of mind tasks.

Keywords: social cognition, theory of mind, appearance–reality, perspective taking

Perception is mostly veridical and informs agents about the way the world is. But sometimes an object or situation appears in a way such that it leads an agent to misconstrue and act upon it in inappropriate ways. For example, when encountering a motionless stick-insect, one may, at first, take it to be a stick and reach out one's arm to grasp it. Only as it suddenly moves does one come to realize what it really is.

A full appreciation of the relationship between appearance and reality requires balancing two complementary facts. On the one hand, one must know that perception is a good (perhaps the best) guide for appropriate actions and beliefs. On the other hand, one must also know that perceptions can be nonveridical and lead to misguided behavior and false assumptions. This is critically important knowledge because it allows one to better identify the sources of one's own as well as others' errors and even predict, in some cases, when these are likely to occur. In line with this Janus-faced character, the terms *appear* and *look like* are highly ambiguous, as they are sometimes used to state what is—probably—the case ("Looks like she's home!") but also figure in contexts in which there is a divergence from reality ("This plastic flower looks real!"; see Austin, 1962; Hansen, & Markman, 2005). Some languages provide different words for objects whose appearance is known to lead to false judgments. In German, for example, *Schein* implies that the way things appear differs from how they are, whereas *Anschein* conveys that one sees no reason to doubt that things are the way they appear to be.

Understanding the relationship between appearance and reality is considered one of the major intellectual achievements in the course of young children's cognitive development. As with children's understanding of beliefs, which becomes manifest in their appreciation of beliefs that contrast with reality, their understanding of the appearance–reality relationship is tested in scenarios in which appearance and reality diverge. In a pioneering study, Flavell, Flavell, and Green (1983) showed preschoolers an object that at a first glance looked like a rock; manual inspection, however, revealed that the object was really a sponge. Children were then asked two questions: (a) what the object looks like and (b) what it really and truly is. The result was that most children between the ages of 4 and 5 answered both questions correctly and judged that the object looked like a rock but was really a sponge. The vast majority of 3-year-olds, in contrast, tended to give "realist responses" and said that the object not only *was* a sponge but also looked like one. One may justify such replies by pointing out that once one has discovered what the object really is, its appearance adjusts accordingly (i.e., the sponge looks more spongy and less rocky once one knows what it is). But note that this cannot explain why children between 4 to 5 years and older knew how to respond to the appearance question; see Perner, 1991.

This pattern of findings has been replicated in a vast number of studies, including many procedural variations. For example, when a glass of milk is placed behind a red color filter, 3-year-olds insist that it looks like milk, not fruit punch (see Flavell, Green, & Flavell, 1986). Not all situations lead children to make the realism error. When an object's property such as color or size is at stake, children mostly answer phenomenistically and judge that the object *is* the way it looks, for example, that the liquid behind the red color filter *is* red, not white (see Taylor & Flavell, 1984). This is in line with findings showing that children between 3 and 4 years of age do not know that an object appears bigger (Pillow & Flavell, 1986) and clearer (Flavell, Flavell, Green, & Wilcox, 1980) when it is close versus distant from an observer. Also by the age of

This article was published Online First October 31, 2011.

Henrike Moll and Michael Tomasello, Department of Developmental and Comparative Psychology, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany.

We would like to thank Birgit Teichmann, Isabelle Gaillande-Mustoe, and Sarah Menzel for assistance with data collection and coding, as well as Roger Mundry for statistical advice. This research project was funded by the Volkswagen Foundation, with a Dilthey Fellowship awarded to the first author.

Correspondence concerning this article should be addressed to Henrike Moll, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany. E-mail: moll@eva.mpg.de

around 4, children begin to grasp that a person may be emotionally touched even though he or she appears indifferent (Harris, Donnelly, Guz, & Pitt-Watson, 1986).

Children's success in the appearance–reality tasks not only temporally coincides with, but also correlates with their performance in other so-called theory of mind tests, such as the standard false belief test (Gopnik & Astington, 1988; Moore, Pure, & Furrow, 1990) and Level 2 visual perspective taking (Flavell, Green, & Flavell, 1986). Three-year-olds' difficulties with the appearance–reality distinction seem robust and profound, as they are not easily remedied by extensive training procedures in which they are playfully taught the meaning of *looks like* and *really is* (Taylor & Hort, 1990). Taken together, these findings support the notion of a major conceptual change taking place between ages 4 and 5 that allows children to understand that objects can be construed in alternative ways (Flavell, 1993; Perner, 1991), including the knowledge that things can look one way but actually be another.

Some investigators have suggested that children's poor performance in the classic tests stems not from any kind of conceptual limitation but from unnecessarily high information-processing demands (Rice, Koinis, Sullivan, Tager-Flusberg, & Winner, 1997) or the odd discourse and linguistic complexity of the tests (Deák, Ray, & Brenneman, 2003; Siegal, 1991; Siegal & Peterson, 1994). Sapp, Lee, and Muir (2000) expected 3-year-olds to demonstrate a full-blown understanding of appearance and reality when a nonverbal response format was used. In their studies, 3.5-year-olds were able to determine which of several objects with different appearances could be placed in front of a camera when someone wanted to take a picture of, for example, a rock (the "rock-sponge"), or which of several objects with different functions could be used to, for example, wipe the table. Hansen and Markman (2005) argued from a discourse-based account that the phrase "looks like" is not always used to note a divergence from reality but also to state what something really is (as in: "What is this?" "Well, what does it look like?"; see also Austin, 1962). Preschoolers performed significantly better in a pragmatically improved version in which it was emphasized, either in the previous discourse or nonlinguistically, that "looks like" in this context implied a contrast to reality. The common tenor of these critical voices is that preschoolers' difficulties with appearance and reality are a mere artifact created by unnecessary linguistic challenges (e.g., Deák, 2006).

We take a different theoretical approach that on the one hand grants 3-year-olds some understanding of appearance and reality but on the other adheres to the idea of a genuine conceptual limitation that does not allow them to fully appreciate the interrelationship between appearance and reality. Using Perner, Stummer, Sprung, and Doherty's (2002) terminology, we claim that children between 4 and 5 learn to confront a deceptive object's appearance with its identity. That is, they are coming to understand that a given object can appear to be one thing but at the same time really be another. Correct answers in the false belief test, Level 2 visual perspective taking, and the alternative naming task (see Doherty & Perner, 1998) are all manifestations of the same basic ability to confront one way of seeing or construing something with another way of viewing or construing the identical object or situation. This view is consistent with that of others who have stressed that preschool children cannot accept different representations or construals of the same thing (e.g., Flavell's, 1988, dual coding hypothesis).

Despite children's difficulties with confronting different perspectives on the same object at the same time, they are surprisingly competent at taking specific perspectives. Unlike a confrontation of two perspectives on the same thing, perspective taking is achieved by determining the referent of another's speech act or goal of his or her action when this is contingent upon a (visual, conceptual, or epistemic) viewpoint that differs in important ways from the child's own. For example, Moll and Meltzoff (2011) found that 3-year-olds know to which of two blue objects an adult is referring as the "green" one (because he or she sees it through a yellow color filter) when the children themselves see both objects in their true, blue color. In the epistemic domain, infants as young as 18 months understand that an agent trying to open a box whose content has been moved either unbeknownst to her or in her full view must, in the first case, be striving for the dislocated content, or, in the second case, be approaching the box for some other reason, such as to place something inside (Buttelmann, Carpenter, & Tomasello, 2009). In these situations, children need not confront two alternative ways of viewing or (epistemically) construing the same state of affairs; they simply have to be able to take or adopt a certain perspective that differs from their own (which they can ignore at that time) and figure out what a person is doing or referring to.

Our conjecture is that the distinction between taking and confronting perspectives can be applied to the appearance–reality situation as well. We therefore conducted two studies. In Study 1, 3-year-olds had to determine which of two objects an adult referred to when she made a request for, for example, "the real" chocolate or for what "looks like" chocolate. In this study, children were not required to confront two (seemingly contradictory) perspectives on a single object; instead, taking the appropriate conceptual perspective on each object individually was sufficient. We therefore predicted that the 3-year-olds should be able to distinguish appearance and reality in this situation. In Study 2, a new sample of 3-year-olds was asked to judge, nonverbally, what a deceptive object looked like and what it really and truly was (e.g., looks like chocolate but really is an eraser). This version of the task forced children to confront or contrast a single object's deceptive appearance with its true identity. We thus predicted that children should not be able to distinguish between appearance and reality in this "confrontational" fashion. If children in the two studies indeed behaved in these ways, it would provide strong evidence that children's struggles in appearance–reality tasks (and perhaps other tasks) is due not to their inability to understand appearance and reality at all but rather to their inability to confront the two perspectives on the same object simultaneously.

Study 1

Method

In this study, children were presented with pairs of objects consisting of, for example, a chocolate bar and a deceptive object that looked like a chocolate bar but was not. After children were shown the objects' functions (and thereby learned about the deceptive object's true identity), they were asked to show which object was the real X and which object looked like X. In order to

do this, children did not have to confront an object's deceptive appearance with its identity, and so we expected 3-year-olds to be successful in this task.

Participants. Twenty-four children (12 boys, 12 girls) of 3 years (M = 38 months, 20 days; range = 35 months, 14 days–39 months, 26 days) participated in this study. They were recruited by telephone from the institute's database of parents who had volunteered to participate with their child in studies on child development. All children were White and from mixed socioeconomic backgrounds. One additional child was tested but excluded from the analyses because she was uncooperative, as she refused to engage with the objects.

Materials. Seven pairs of objects were used as stimuli. Each pair consisted of a regular exemplar of a certain category (nondeceptive object) and another item that appeared to be an exemplar of the same category but, as per its function, was something else (deceptive object). The first pair, which was always used during a demonstration phase preceding the test, consisted of a syringe and a pen that looked like a syringe. For the test phase, the following six pairs of objects were used: a children's book and a box that looked like a book, a tube of toothpaste and a pen that looked like a tube of toothpaste, a pen and an eraser that looked like a pen, a lemon, and a kitchen alarm bell that looked like a lemon, a transparent box with real stones and a transparent bag with candy that looked like stones, and a small chocolate bar and an eraser that looked like a chocolate bar (see Figure 1a). All objects were 1–12 cm wide, 5–13 cm long, and 0.5–8 cm high. A previously conducted test with a separate sample (n = 17) of 3-year-olds revealed that the vast majority of children in the participants' age range was indeed "tricked" by the appearances of the deceptive objects and, when first seeing them, judged, for example, that the eraser-chocolate was chocolate (one child identified the stone-candy as candy, but this was the only case of a correct identification of a deceptive object). The children thus not only saw a vague similarity or resemblance between the deceptive objects and what they appeared to be but actually tended to misconstrue them accordingly. Props were used to demonstrate the functions of some of the objects. This was a bowl with water (to fill the syringe), a sheet of paper with pencil lines (used to demonstrate the erasers and pens), a toothbrush (for the toothpaste), and a small blunt toy knife made out of plastic (used to pretend to cut the stone-candy and a piece from the lemon, which were both precut). A white tray (30 × 40 × 1 cm) was used from which children chose the objects during the test phase.

Design. The study involved one experimenter (E). For each pair of test objects, she made one request for the deceptive object (Appearance Request) and one request for the nondeceptive object (Reality Request). There were thus six requests of each type and a total of 12 requests. The order of requests (Appearance Request first vs. Reality Request first), the order in which children were familiarized with each object within a pair (deceptive object first vs. nondeceptive object first), and the spatial arrangement of the objects in the tray at test (deceptive object left vs. deceptive object right) were counterbalanced. Every child received the six pairs of test objects in a different order.

Procedure. The child, the parent, and E briefly played in a greeting lobby outside of the testing room. The parent was instructed not to intervene during the experiment and to sit silently throughout the session. They then entered the testing room (436 × 441 cm), where each took their predefined seats. E and the child sat facing each other at a small table (59 cm wide, 50 cm long, 50 cm high), and the parent was seated 80 cm behind the child. On the table were a bowl of water and the demonstration objects (the syringe and the pen that looked like a syringe).

Demonstration phase. E took the syringe and explained that it could be used to splash water. She pulled water into the syringe and splashed it back into the bowl. She explained, "This is a real, true syringe." She gave the syringe to the child and encouraged him or her to use it. When the child was finished, E held up the pen and explained that this "only looks like a syringe" and that it could not be used for splashing water. E said that it can be used for drawing instead and drew some lines on a piece of paper. She then passed the pen on to the child and encouraged him or her to draw as well. When the child was finished, E picked up the pen and repeated, that "This only looks like a syringe—it isn't one, really." She then held up the syringe, saying, "*This* is a real, true syringe." At the end of this demonstration, the syringe and the pen were placed next to each other on the tray, which was slid toward the child. Looking into the child's face throughout the utterance of her request, E asked the child to show "the real, true syringe." If the child chose the wrong object, E corrected him or her by holding up the target object, saying "*This* is the real, true syringe." A negative response was again followed by a correction and repetition of the request. E then moved on to the next question and asked for the

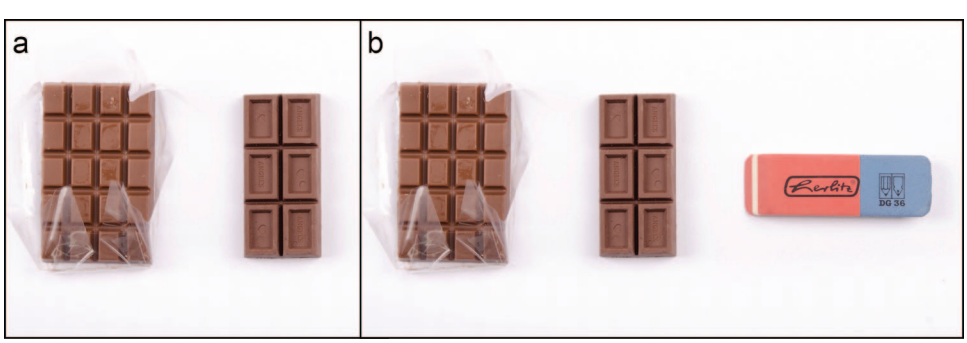

Figure 1. A sample of the objects used in Studies 1 and 2. 1a: One of the pairs of test objects from Study 1: nondeceptive object on the left, deceptive object on the right. 1b: One of the sets of test objects from Study 2: deceptive object in the middle, appearance-object on the left, and reality-object on the right.

object that "only looks like a syringe." The same correction procedure and repetition of the request were applied, if needed. Finally, the training objects and tray were removed from the table.

Familiarization with the test objects. After this demonstration phase, *E* placed the first pair of test objects on the table. She then showed how the two objects were to be used (in counterbalanced order). For example, she took the chocolate-eraser and exclaimed that "one can erase lines with this" and erased some pencil lines on a sheet of paper. She passed the object to the child, saying, "Now you can erase some lines!" The child then used the object herself. *E* used only functional descriptions for the objects, for example, "One can draw/erase/look at pictures/eat/open (with) this," but never labeled them. Finally, the two objects were placed next to each other on the tray at a distance of 5–10 cm from each other (in counterbalanced spatial order).

Test phase. Looking into the child's face throughout the response phase, *E* made a request for the deceptive object (Appearance Request) or the nondeceptive object (Reality Request). When asking for the deceptive object, she exclaimed, "Please show me what only looks like chocolate!" When asking for the nondeceptive object, she said, "Please show me the real, true chocolate!" Children responded by pointing at or touching the objects. After they responded, *E* said "Okay!" and pulled the tray back toward her side of the table. She then uttered the next request, that is, the Reality Request if the Appearance Request had been made first, and vice versa. After the child responded, *E* again said, "Okay!," removed the objects from the table, and brought out the second pair of test objects. The procedure was repeated in the exact same manner for the remaining pairs.

Scoring and reliability. The videotaped responses were scored by *E*. She recorded for each trial whether the child chose the deceptive object, the nondeceptive object, or no response was made. If both objects were pointed at or touched one after the other, the first response was scored, unless the child made verbally clear that she aimed to "self-correct"—in which case the second choice was scored. Choices of the object that matched the content of the request were coded as *correct* (1), and choices of the object that did not match the request were coded as *incorrect* (0). If a child pointed at or touched neither object or both of them simultaneously without later disambiguating his or her response (by pointing at or touching a single object), his or her choice was scored as *incorrect* (0). To assess interobserver reliability, an independent research assistant who was ignorant of the content of the request (the sound was turned off during the request) coded a randomly selected sample of 25% of the children. There was disagreement between the observers on one of the responses, which was easily resolved. Cohen's kappa was .99. Two trials were disregarded because of uncooperative behavior; therefore mean percentages are reported.

Results and Discussion

Figure 2 shows the mean percentage of correct choices by children as a function of request type. The children's gender and the order of requests had no effect on the choices, $p = .15$ and $p = .57$, respectively. On average, children chose the correct object in 62% of the cases when an Appearance Request was made and in 69% when a Reality Request was made. As shown by one-sample *t* tests, the level of correct responses significantly exceeded chance (set at 50%), both when the children received an Appearance Request, $t(23) = 2.21$, $p < .04$, and when they received a Reality Request, $t(23) = 3.93$, $p < .01$.

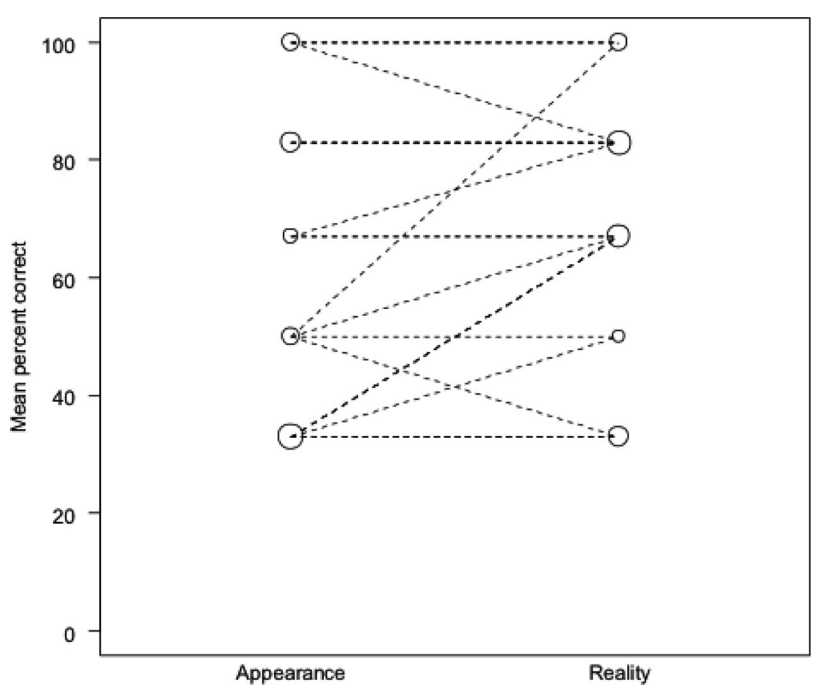

Figure 2. Mean percentage of correct responses as a function of question type obtained by children in Study 1. Areas of circles are proportionate to the number of children who achieved a given percentage of correct choices.

To compare the difficulty of the Appearance Request and the Reality Request, we used each child as his or her own control and compared the number of successful trials for each of the two request types using a paired-sample t test. Children gave significantly more correct responses when they were presented with a Reality Request than when they received an Appearance Request, $t(23) = 2.09$, $p < .05$.

The 3-year-old children in this study were presented with a verbal request for either a deceptive object or its "real" counterpart. Children could choose between an object that belonged to a certain category (e.g., chocolate) and another one that, at first sight, appeared to belong to the same category but upon closer scrutiny turned out to be something else (e.g., an eraser). As predicted, the 3-year-olds mainly responded appropriately to the requests. When presented with an Appearance Request, they chose the deceptive object significantly above chance. Likewise, when presented with a Reality Request, they reliably chose the nondeceptive object at a level exceeding chance.

In their studies with 3.5-year-olds, Sapp et al. (2000) aimed to show that children younger than 4–5 years can distinguish between appearance and reality when a nonverbal response format is applied. Importantly however, in three out of four experiments, the candidate objects differed not only with respect to their identity (operationalized via different functions), but their appearance as well. When asked for something that could be placed in front of a camera if they wanted to take a picture of a rock, children had to choose between the rock-sponge and things that looked nothing like a rock, such as a cup and an apple. Note that no understanding of the difference between "real" versus "looks like" was necessary to select the correct object. Not surprisingly, children performed at ceiling. In the final experiment, children were given the choice between a real rock and the rock-sponge, but the requests did not differ along the real–apparent dimension. In the reality request, the adult simply requested something with which to wipe off dirt. In this situation, a basic knowledge of the objects' affordances (Gibson, 1977) was sufficient for a correct response. The two candidate objects in our study, in contrast, differed only with respect to their (mis)match between appearance and identity, and the requests were articulated along this very dimension.

It must be noted that children's performance level was far from excellent—especially when presented with an Appearance Request. There may be different reasons for children's moderate performance level and their difficulties with the Appearance Request specifically. First, before children learn about the complex relationship between appearance and reality, they have to build up a solid understanding of what things are, what their functions are, and how they are called. Possibly by the age of 3, many children have only started to enter the discourse about objects looking like one thing but really being another. Second, real chocolate "looks like" chocolate, too, and so choosing it as a response to the Appearance Request may be seen as a perfectly appropriate response. In most contexts, however, referring to something that is mutually known to be chocolate as the thing that "looks like chocolate" when there has been neither doubt nor denial that it *is* chocolate is pragmatically odd (see Grice, 1961). Even if children are not familiar at any explicit level with such pragmatics, confusion about which particular use of "looks like" was invoked in the test was highly unlikely, given the circumstances and context (deceptive objects paired with nondeceptive ones) and given that the terminology was explained to children in the demonstration phase. Nonetheless, to prevent any potential for confusion, we had the adult request what "*only* looks like X." While both objects, deceptive and nondeceptive, may have been said to "look like X" in some sense, the deceptive object alone could be correctly described as the one that "*only* looks like X."

A better account of the residual realism bias found in this study is to see it as resulting from a "pull of the real" similar to the one found in other theory of mind tasks. The reason possibly lies in the combination of parts of the content of the request and children's knowledge. The Appearance Request contained the word *chocolate*, and children knew which object in front of them was chocolate. Nonetheless, this was precisely the object they were required *not* to choose. This situation is similar to the false belief test in which children hear a question that contains a reference to an object's whereabouts or identity that is known by the child. Getting the answer right requires inhibiting this knowledge (e.g., Sabbagh, Xu, Carlson, Moses, & Lee, 2006) and paying close attention to the crucial part of the question that specifies that the children should, in fact, *not* report where or what the object is but how the misinformed agent conceives of it. Despite these challenges, most of the 3-year-olds in this study showed some understanding of the difference between deceptive and nondeceptive objects and how each of them was captured conceptually or linguistically.

Our conjecture is that many 3-year-olds were competent when facing this new appearance–reality task because it allowed them to determine which object was appropriately construed as the "real, true X" and what "only looks like X." Because these terms captured a particular way of construing an object, we argued that children could take such conceptual perspectives and determine the referent. However, we predicted that 3-year-olds would not be successful in a similar task that afforded a confrontation of these two conceptual perspectives on the very same thing (a deceptive object). The rationale was that children at this age cannot yet explicitly judge that the very same object looks like one thing but really is another—even when the same selection measure is applied. To test this, a second study was conducted with a task that could not be solved by distinguishing real from apparent, for example, chocolate bars. Instead, the two conceptual perspectives, (a) what an object appears to be according to its phenomenology and (b) what it is according to its function, had to be confronted on one and the same object.

Study 2

Method

As in Study 1, the 3-year-old children in this study were presented with verbal questions in a forced-choice task. They were asked to indicate what a deceptive object (e.g., a chocolate-eraser) looked like and what it really was by pointing to either (a) an exemplar of the category to which the object *appeared to* belong (e.g., a chocolate bar) or (b) an exemplar of the category to which the object *really* belonged (e.g., a regular eraser). Children thus had to nonverbally "confront" the deceptive object's appearance with its true identity. The prediction was that, unlike children of the same age in Study 1, the 3-year-olds in this study would perform poorly because of their inability to confront two conceptual perspectives on an identical object.

Participants. Participants in this study were 24 (12 boys, 12 girls) 3-year-olds (M = 38 months, range = 34 months, 10 days–39 months, 26 days). No additional children were tested but excluded from the analyses. The subjects were taken from the same database of parents who had volunteered to participate in studies on child development as subjects in Study 1.

Materials. For the demonstration phase, the same syringe and pen that looked like a syringe were used as in Study 1. For the test phase, six sets of three objects were used. Each set consisted of one deceptive object (all the deceptive objects were identical to those used in Study 1), one exemplar of the category to which the deceptive object really belonged (reality-object), and one exemplar of the category to which the deceptive object appeared to belong (appearance-object). The following six sets of objects were used: (a) a book, a box that looked like a book, and a regular box; (b) a tube of toothpaste, a pen that looked like a tube of toothpaste, and a regular pen; (c) a pen, an eraser that looked like a pen, and a regular eraser; (d) a lemon, a kitchen alarm bell that looked like a lemon, and a regular bell; (e) a transparent box with stones, a transparent bag with candy that looked like stones, and regular candy; (f) a small chocolate bar, an eraser that looked like a chocolate bar, and a regular eraser. As Figure 1b shows, the deceptive object strongly resembled the appearance-object but was visually dissimilar to the reality-object. All objects were 0.5–10 cm wide, 4–15.5 cm long, and 0.5–8 cm high.

Design. For each set of test objects, children received two questions: one in which they had to indicate what the deceptive object looked like (Appearance Question), and one in which they had to indicate what it really was (Reality Question). There were thus six questions of each type and a total of 12 questions. The order of question type (Appearance Question first or second), the order in which children were familiarized with each of the three objects in a set, the spatial arrangement of the two objects to choose from at test in the tray (reality-object left or right), as well as the order in which *E* pointed at the two candidate objects while asking her question (appearance-object first vs. reality-object first), were counterbalanced. Each child received a different order of the six sets of test objects.

Procedure. The beginning of the procedure was identical to that in Study 1, including the demonstration phase. The children were familiarized with the test objects in the same manner as in Study 1—the only difference being that three objects (deceptive object, appearance-object, and reality-object) were shown to them. As in Study 1, *E* never labeled these objects and used only functional descriptions such as, "One can draw with this."

Test phase. After children were familiarized with the three objects, *E* placed the appearance-object (e.g., the chocolate) and the reality-object (e.g., the regular eraser) next to each other on the tray at a distance of approximately 5–10 cm (in counterbalanced spatial order). These two were the candidate objects between which children had to choose at test. *E* then held up the deceptive object (e.g., the chocolate-eraser), pointed at it with the index finger of her free hand, made eye contact with the child, and then posed the question. When asking an Appearance Question, she asked, "What does this look like: like *this* [pointing at first candidate object] or like *this* [pointing at second candidate object]?" When asking a Reality Question, she asked, "What is this really? One of these [pointing at first candidate object] or one of these [pointing at second candidate object]?" Children responded by pointing at or touching one of the two objects. As soon as a child made her choice, *E* said, "Okay!" and pulled the tray back toward her side of the table. She then uttered the next question. This was an Appearance Question if the Reality Question had already been posed and vice versa. After the child responded, *E* again said "Okay!," removed the objects from the table, and brought out the second set of objects. This procedure was repeated in the exact same manner for the remaining sets of objects.

Scoring and reliability. The videotaped responses were scored by *E*. She recorded for each trial whether the child chose the reality-object or the appearance-object. The scoring procedure was identical to that used in Study 1. To assess interobserver reliability, an independent research assistant who was ignorant of the content of the question (the sound was turned off during the question) coded a randomly selected sample of 25% of the children. There was a disagreement between the two observers about one of the responses, which was easily resolved. Cohen's kappa was .97. One child was uncooperative during two trials which were therefore disregarded. Mean percentages are reported for this reason.

Results and Discussion

Figure 3 shows the mean percentage of correct choices by children as a function of question type. Gender and order of questions had no effect on the children's performance, $p = .17$ and $p = .31$, respectively. On average, children chose the appearance-object in 64% of the cases when an Appearance Question was asked, and they chose the reality-object in 34% of the cases when a Reality Question was asked. As shown by one-sample t tests, their level of correct responses was significantly above chance (set at 50%) when an Appearance Question was asked, $t(23) = 2.54$, $p < .02$ but significantly *below* chance when a Reality Question was asked, $t(23) = 4.00$, $p < .01$.

To compare the difficulty of the two types of question, we used each child as his or her own control and compared the number of successful trials for each of the two question types using a paired-sample t test. Children produced significantly more correct responses to Appearance Questions than to Reality Questions, $t(23) = 3.69$, $p < .01$.

In this study, 3-year-old children were not able to respond correctly to both questions of a nonverbal appearance–reality task. While the vast majority of children correctly identified what a deceptive object (e.g., a chocolate-eraser) looked like, they failed to indicate what it really was. Instead of choosing a typical exemplar of the category to which the deceptive object belonged as per its function, the children significantly chose the object that matched its appearance. In other words, they did not succeed in confronting a deceptive object's appearance with its true identity but indicated how the object looked irrespective of what they were asked.

This result seems to contrast with the one Rice and colleagues (1997) obtained using a version with reduced information-processing demands. In their study, children between 3 and 3.5 years benefited from seeing a rock and a sponge on the table as they answered the standard appearance–reality questions. The goal of placing the two representations in front of the children was to make it unnecessary for them to hold the two conflicting representations in mind and compare them. We cannot say exactly why children benefited from the presence of the objects, given that the

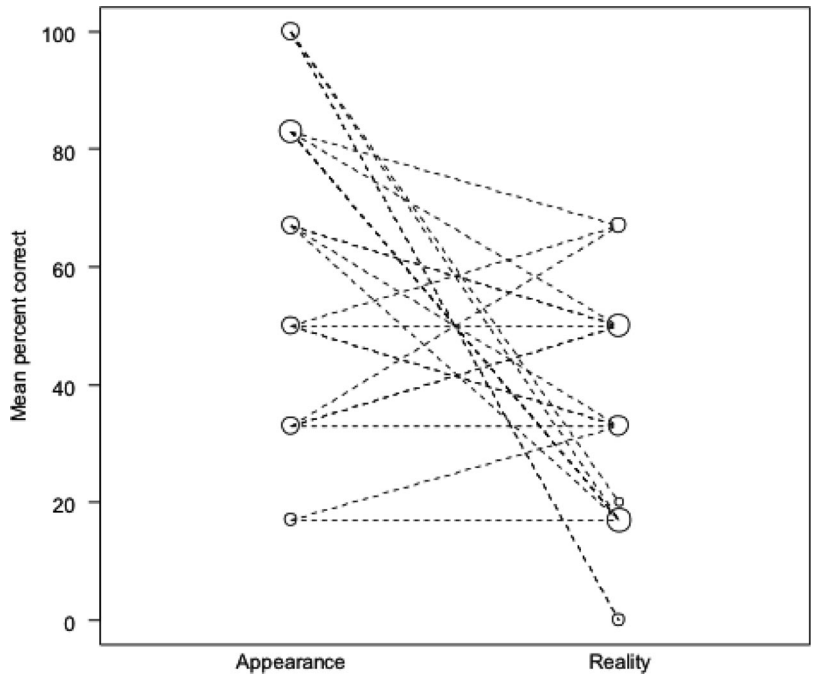

Figure 3. Mean percentage of correct responses as a function of question type obtained by children in Study 2. Areas of circles are proportionate to the number of children who achieved a given percentage of correct choices.

task still required the confrontation of two perspectives on the very same object. But it seems unlikely that remembering the words posed a problem in the standard test: They knew both words (as evidenced by their reactions when first shown the object) but then systematically give the wrong answer to one particular question. The present study adds to the evidence that uttering or remembering the words is not the reason why children struggle. Future research needs to explore why children in Rice and colleagues' (1997) test nonetheless showed an improved performance.

Unlike previous appearance–reality tests in which children had to judge about an object's identity rather than a property such as color or size, the 3-year-olds in this study mostly made the phenomenism error—judging that the deceptive object not only looks like but really is what it appears to be. We believe that this is so because the adult did not provide the labels for the two alternative construals (*eraser* and *chocolate*), as is done in the traditional test version ("What does this look like, a *sponge* or a *rock*?"). In this classic variant, the children respond on the basis of what they know the object to be (a rock) and therefore engage in the realism error. This adds to converging evidence that the particular way in which a task is conceptually framed determines which of the two errors children will tend to make—even when the physical set-up and procedure are identical (Taylor & Flavell, 1984). It also lends support to the point made in the discussion of Study 1, that providing the labels is what gears children toward realist responses. Presenting children with the two alternative conceptualizations *rock* and *sponge* biases them toward reality because it brings to their mind what the object in fact is and thus ought to be called. In this study, in which the terms were not provided, realism gave way to phenomenism because visual similarities gain dominance in the absence of concepts. But irrespective of the type of error children make, the results support the view that 3-year-olds have difficulties when they have to judge both what a deceptive object appears to be or is according to its phenomenology and what it is according to its function.

General Discussion

Developmental researchers have traditionally argued that the distinction between appearance and reality remains completely opaque to children until the *threshold age* has been reached between 4 and 5 years (Flavell, 1986; Perner, 1991). More recently, the pendulum has swung the other way, with the argument that younger children do in fact possess a full appreciation of this distinction but are hindered in making it manifest in the classic tests because of superfluous task demands, such as unnecessary verbal requirements (Deák, 2006; Sapp et al., 2000), misleading discourse (Hansen & Markman, 2005), or processing demands (Rice et al., 1997).

The results of the present studies suggest that neither of the two camps provides a fully accurate account of 3-year-olds' abilities in this cognitive domain. Of the two novel appearance–reality tests that 3-year-olds were administered, they showed competence in one (Study 1) but not the other (Study 2)—despite the superficial similarities of the two test variants, such as having to choose between one of two objects in response to an adult's verbal request or question. A more nuanced view of young children's understanding of appearance and reality is therefore needed. We propose that the distinction between *taking* and *confronting* (Perner et al., 2002) perspectives provides a useful conceptual framework that allows for a better description of what 3-year-olds have and have not yet come to understand in this social-cognitive area.

As is shown by their mainly correct responses in Study 1, a significant number of 3-year-olds know which of two objects is a real versus apparent exemplar of a certain kind of object. This is the youngest age for which a basic understanding of appearance and reality has been unequivocally demonstrated. Yet these children did not have to confront two construals of the same object and reflect upon the fact that the deceptive object, while having the appearance of chocolate, was in fact an eraser. Instead, they could solve the task by determining the referent of the adult request for either object: Only the chocolate could be referred to as "real, true chocolate," whereas only the eraser could be referred to as "what only looks like" chocolate. That most children at age 3 are able to make this determination is in line with the view that they can "take" perspectives—whether these are conceptual or perceptual perspectives. In a recent Level 2 perspective taking test, 3-year-olds had no problem determining to which of two objects an adult referred as the green and the blue one—even though the children themselves saw both objects as blue. They were able to take the adult's perspective on the two objects and disambiguate her request accordingly (Moll & Meltzoff, 2011). In that study, children did not have to become aware that one of the two objects was seen in two different ways.

In Study 2, however, children were unable to distinguish between appearance and reality. Determining the referent of a request was not sufficient to pass this variant of the test. Instead, children had to make an explicit (though nonverbal) judgment about how a deceptive object was construed according to its phenomenology and according to its identity. They had to confront two perspectives, namely, what an object appeared to be at first sight and what it turned out to be upon closer scrutiny. The two options to choose between were not simple targets of an adult request, as in Study 1. Rather, the child had to use them as quasilinguistic items that each represented one of the two ways (i.e., perspectives) in which the deceptive object could be construed. The same logic was applied in a test on children's abilities to confront visual perspectives in which children had to indicate how they and an adult each saw a given object by pointing to either of two color samples (in blue and green) that represented the two different ways of seeing the same thing. Children younger than 4.5 years were not able to solve this task (Moll, Meltzoff, Merzsch, & Tomasello, 2011). Again, determining which object an adult is talking about or referring to—which is sufficient for taking perspectives—is not possible. The candidate objects are used quasilinguistically to judge how a given object is seen from a certain perspective.

Taken together, these findings show that the problem of the classic appearance–reality tasks is not, as some have argued (Sapp et al., 2000), that they demand a verbal response—at least not as long as children are familiar with the relevant vocabulary, such as "looks like" and "really is." Whether they reply by uttering the words or by pointing to objects that represent these terms makes no difference. The difficulties that 3-year-olds are facing in these tests reveal not an articulatory but a conceptual limitation, which becomes manifest both in the standard tests as well as in alternative variants that preserve their cognitive architecture. The defining feature of these tests is that children have to make explicit judgments, verbal or nonverbal, about how an object is construed from a particular perspective that directly confronts another way of seeing or construing the same object or state of affairs of which the child is simultaneously aware (see also Frye, Zelazo, & Burack, 1998).

This also means that young children's inability to confront perspectives becomes apparent only under specific circumstances but passes unnoticed in many of the regular everyday interactions, which are mainly grounded in perspective taking. For example, even infants understand that an adult asking them for a piece of food may want broccoli, not crackers (even though the infants themselves have the opposite preference, Repacholi & Gopnik, 1997), or that a person striving for an empty box wants to retrieve an object whose removal she failed to witness (Buttelmann et al., 2009)—which is why something like an implicit theory of mind is strongly debated (e.g., Clements & Perner, 1994; Onishi & Baillargeon, 2005; Southgate, Senju, & Csibra, 2007; Surian, Caldi, & Sperber, 2007). Young children and even infants can engage in pretend play and act as if an object were something it is actually not. But again, no confrontation of what the object really is with what it represents during pretense is necessary (see Perner, 1991). In the linguistic domain, children will not find anything wrong with their parent calling the neighbor's pet a *dog* on one occasion and a *puppy* on another (see Doherty, 2000), indeed, they themselves most likely make use of both terms in different discourse episodes. What is not yet in young children's cognitive repertoire is an explicit acknowledgment of the possibility of alternative views or construals of the same object.

A final note needs to be made on the problems that we see with treating appearance and reality as oppositions (see Austin, 1962; Hacker, 1987; Nudds, 2011). In appearance–reality tests, reality is operationalized via an object's function. But an object's appearance is just as real as its function. This becomes especially clear when the focus of interest shifts away from what kind of thing an object is to its gestalt. For example, a stick when held under water certainly does have the gestalt of a bent rod. With this focus in mind, a ball-shaped candle *is* a ball, and a bell-shaped piece of marzipan *is* a bell. Thus, when facing several differently shaped pieces of marzipan, referring to the "bell" would be perfectly reasonable. It should also be kept in mind that what prompts a revision of one's initial assumptions about a "deceptive" object is yet another, second look at it or the use of other sense modalities. It is thus again an appearance that leads us to correct our previous misconstrual. In line with these remarks, the term *looks like* is often used in situations that do not involve any distance from reality at all, as when we apply this term comparatively to state a resemblance with some other object that it is very unlikely or even impossible to get confused with (e.g., "Paul looks like his sister") or when we use it evidentially to say what we think is the case (e.g., "Looks like Paul is home" when we see that the lights in his living room are on). Future research should take account of these issues that have been neglected in developmental research.

The current results thus specify the nature of young children's difficulties in the appearance–reality task in more detail than have previous studies, and they suggest, in addition, a larger theoretical framework that may help to explain children's similar difficulties in related tasks such as false belief or alternative naming. If extended fruitfully to these other tasks, the current framework may help us to better understand the important cognitive transition that occurs between around 4 and 5 years of age.

References

Austin, J. L. (1962). *Sense and sensibilia.* Oxford, England: Oxford University Press.

Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition, 112,* 337–342. doi:10.1016/j.cognition.2009.05.006

Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development, 9,* 377–395. doi:10.1016/0885-2014(94)90012-4

Deák, G. O. (2006). Do children really confuse appearance and reality? *Trends in Cognitive Sciences, 10,* 546–550. doi:10.1016/j.tics.2006.09.012

Deák, G. O., Ray, S. D., & Brenneman, K. (2003). Children's perseverative appearance–reality errors are related to emerging language skills. *Child Development, 74,* 944–964. doi:10.1111/1467-8624.00578

Doherty, M. (2000). Children's understanding of homonymy: Metalinguistic awareness and false belief. *Journal of Child Language, 27,* 367–392. doi:10.1017/S0305000900004153

Doherty, M., & Perner, J. (1998). Metalinguistic awareness and theory of mind: Just two words for the same thing? *Cognitive Development, 13,* 279–305. doi:10.1016/S0885-2014(98)90012-0

Flavell, J. H. (1986). The development of children's knowledge about the appearance–reality distinction. *American Psychologist, 41,* 418–425. doi:10.1037/0003-066X.41.4.418

Flavell, J. H. (1988). The development of children's knowledge about the mind: From cognitive connections to mental representations. In J. W. Astington, P. L. Harris, & D. R. Olson (Eds.), *Developing theories of mind* (pp. 244–267). Cambridge, England: Cambridge University Press.

Flavell, J. H. (1993). The development of children's understanding of false belief and the appearance–reality distinction. *International Journal of Psychology, 28,* 595–604. doi:10.1080/00207599308246944

Flavell, J. H., Flavell, E. R., & Green, F. L. (1983). Development of the appearance–reality distinction. *Cognitive Psychology, 15,* 95–120. doi:10.1016/0010-0285(83)90005-1

Flavell, J. H., Flavell, E. R., Green, F. L., & Wilcox, S. A. (1980). Young children's knowledge about visual perception: Effect of observer's distance from target on perceptual clarity of target. *Developmental Psychology, 16,* 10–12. doi:10.1037/0012-1649.16.1.10

Flavell, J. H., Green, F. L., & Flavell, E. R. (1986). Development of knowledge about the appearance–reality distinction. *Monographs of the Society for Research in Child Development, 51*(1, Serial No. 212). doi:10.2307/1165866

Frye, D., Zelazo, P. D., & Burack, J. A. (1998). Cognitive complexity and control I: Theory of mind in typical and atypical development. *Current Directions in Psychological Science, 7,* 116–121. doi:10.1111/1467-8721.ep10774754

Gibson, J. J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing: Toward an ecological psychology* (pp. 67–82). Hillsdale, NJ: Erlbaum.

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance–reality distinction. *Child Development, 59,* 26–37. doi:10.2307/1130386

Grice, H. P. (1961). The causal theory of perception. *Proceedings of the Aristotelian Society* [Supplement no. 35], 121–152.

Hacker, P. M. S. (1987). *Appearance and reality. A philosophical investigation into perception and perceptual qualities.* Oxford, England: Blackwell.

Hansen, M. B., & Markman, E. M. (2005). Appearance questions can be misleading: A discourse-based account of the appearance–reality problem. *Cognitive Psychology, 50,* 233–263. doi:10.1016/j.cogpsych.2004.09.001

Harris, P. L., Donnelly, K., Guz, G. R., & Pitt-Watson, R. (1986). Children's understanding of the distinction between real and apparent emotion. *Child Development, 57,* 895–909. doi:10.2307/1130366

Moll, H., & Meltzoff, A. N. (2011). How does it look? Level 2 perspective-taking at 36 months. *Child Development, 82,* 661–673. doi:10.1111/j.1467-8624.2010.01571.x

Moll, H., Meltzoff, A. N., Merzsch, K., & Tomasello, M. (2011). *Taking versus confronting visual perspectives in preschool children.* Manuscript submitted for publication.

Moore, C., Pure, K., & Furrow, P. (1990). Children's understanding of the modal expression of certainty and uncertainty and its relation to the development of a representational theory of mind. *Child Development, 61,* 722–730. doi:10.2307/1130957

Nudds, M. (2011). Children's understanding of perceptual appearances. In J. Roessler, H. Lerman, & N. Eilan (Eds.), *Perception, causation, and objectivity* (pp. 264–285). Oxford, England: Oxford University Press.

Onishi, K., & Baillargeon, R. (2005, April 8). Do 15-month-old infants understand false belief? *Science, 308,* 255–258. doi:10.1126/science.1107621

Perner, J. (1991). *Understanding the representational mind.* Cambridge, MA: MIT Press.

Perner, J., Stummer, S., Sprung, M., & Doherty, M. (2002). Theory of mind finds its Piagetian perspective: Why alternative naming comes with understanding belief. *Cognitive Development, 17,* 1451–1472. doi:10.1016/S0885-2014(02)00127-2

Pillow, B. H., & Flavell, J. H. (1986). Young children's knowledge about visual perception: Projective size and shape. *Child Development, 57,* 125–135. doi:10.2307/1130644

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology, 33,* 12–21. doi:10.1037/0012-1649.33.1.12

Rice, C., Koinis, D., Sullivan, K., Tager-Flusberg, H., & Winner, E. (1997). When 3-year-olds pass the appearance–reality test. *Developmental Psychology, 33,* 54–61. doi:10.1037/0012-1649.33.1.54

Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind: A comparison of Chinese and U.S. preschoolers. *Psychological Science, 17,* 74–81. doi:10.1111/j.1467-9280.2005.01667.x

Sapp, F., Lee, K., & Muir, D. (2000). Three-year-olds' difficulty with the appearance–reality distinction: Is it real or is it apparent? *Developmental Psychology, 36,* 547–560. doi:10.1037/0012-1649.36.5.547

Siegal, M. (1991). *Knowing children: Experiments in conversation and cognition.* Hove, England: Erlbaum.

Siegal, M., & Peterson, C. C. (1994). Children's theory of mind and the conversational territory of cognitive development. In C. Lewis & P. Mitchell (Eds.), *Children's early understanding of mind: Origins and development* (pp. 427–455). Hove, England: Erlbaum.

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by two-year-olds. *Psychological Science, 18,* 587–592. doi:10.1111/j.1467-9280.2007.01944.x

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science, 18,* 580–586. doi:10.1111/j.1467-9280.2007.01943.x

Taylor, M., & Flavell, J. H. (1984). Seeing and believing: Children's understanding of the distinction between appearance and reality. *Child Development, 55,* 1710–1720. doi:10.2307/1129918

Taylor, M., & Hort, B. (1990). Can children be trained in making the distinction between appearance and reality? *Cognitive Development, 5,* 89–99. doi:10.1016/0885-2014(90)90014-K

Received February 8, 2011
Revision received September 8, 2011
Accepted September 14, 2011 ■