

Research Article

Appearances Are Deceiving: Observing the World as It Looks and How It Really Is—Theory of Mind Performances Investigated in 3-, 4-, and 5-Year-Old Children

Lucia Bigozzi, Alessandra Di Cosimo, and Giulia Vettori

Department of Education and Psychology, University of Florence, Florence, Italy

Correspondence should be addressed to Giulia Vettori; giulia.vettori@unifi.it

Received 6 July 2016; Revised 24 September 2016; Accepted 7 November 2016

Academic Editor: Randal X. Moldrich

Copyright © 2016 Lucia Bigozzi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Appearance-reality (AR) distinction understanding in preschoolers is worth of further consideration. This also goes for its relationship with false-belief (FB) understanding. This study helped fill these gaps by assessing 3-, 4-, and 5-year-old children's performances on an appearance-reality distinction task and by investigating relationships with unexpected location, deceptive content, and deception comprehension task performances. 91 preschoolers participated in this study divided into 3 groups: (1) 37 children, *M*-age 3.4 years; (2) 23 children, *M*-age 4.5 years; (3) 31 children, *M*-age 5.4 years. A developmental trend was found where appearance-reality distinction understanding was significantly influenced by age. If wrong answers were particularly high by 3-year-old children, they greatly decreased by 4- and 5-year-old children. 3-year-old children also tended to fail in FB tasks; instead 4- and 5-year-old children performed AR tasks better than FB tasks. Theoretical and practical implications were discussed.

1. Theory of Mind: A Complex Construct

Since Premack and Woodruff [1], “theory of mind” (ToM) has been one of the major fields of research in child development. Recognized as a multifaceted sociocognitive process [2], it shows a deep involvement with both cognitive and social functioning [3, 4]. Children's ToM defines both awareness of their own mental states (i.e., thought, decision, knowledge, and belief) and the fact that people may have different representations of the world and act on the basis of them [1]. So, this ability allows one to explain and predict people's behavior [5–7]. The key aspects of ToM are that children recognize other people as psychological beings [8] and distinguish internal from external world [9]. The construct of ToM comprises different components: from young children's ability to speak of their own and others' mental states and to lie [10], to the preschoolers' use of language of mind [11] and ability to comprehend false beliefs [12, 13], and to appearance-reality distinction [14–16]. With this study, we explored appearance-reality distinction understanding in preschoolers and investigated relationships with false-belief understanding.

2. Appearance-Reality Distinction and False Beliefs

Researchers interested in studying ToM in preschoolers focus on the ability to understand that people will act in accordance with their beliefs about reality, even if those beliefs are false [17] and do so, especially, using false-belief tasks, like the unexpected location [18] and deceptive contents [16, 19]. In a typical unexpected location task children are shown a scenario in which a story character places a desirable object (such as a candy or a ball) in a particular location before leaving the scene [18]. Then, another character transfers the object to a different location. The child is then asked to predict where the first character will look for his/her object when he/she comes back. To attribute a different representation of reality to the first character that will influence his/her behavior means mastering a false belief through the recursive process “I think that you think.” The deceptive content (called also “Smarties”) task included two direct questions to investigate both the child's false belief and the others' false belief [19]. The procedure includes using a Smarties tube filled with crayons. The researcher first asked the child what

he/she thought was in there and generally the child responded by saying Smarties, candies, or similar. At this point, the researcher showed that the true content was crayons and then put the crayons back into the tube. The child is asked two questions to investigate his/her own *“What did you first think was inside the tube?”* and the others false belief *“What will your friend think is inside the tube?”*

A less investigated component of ToM in preschool children is the appearance-reality distinction such as understanding that mental states can differ from reality states [6, 14]. This latest concept highlights the fact that theory of mind also means being aware of the fact that our mental world could differ from the physical one and actively try to interpret and reason for the causes and consequences of those possible differences [21]. An experimental paradigm to investigate children’s appearance-reality understanding was introduced and developed by Flavell et al. [14]. In this experiment, the researcher shows the participant a sponge that looks like a stone. Then, the researcher asks the participant if the object looks like a stone or a sponge and whether the same object is actually a stone or a sponge. The correct answers presuppose the ability to distinguish the appearance (as a representation believed to be true) from the exact representation of reality. At the same time, it implies the ability to handle the simultaneous presence of two different representations of the same object. Before reaching this awareness, children tend to believe that their perceptions of the world are accurate reflections of its actual properties and that others will therefore perceive the world as they do, so they are egocentrically biased [22]. Children who understand the distinction between appearance and reality have an awareness of the real nature of the object; they distinguish their representation of reality from that held by others, and they predict that others could be deceived by it. This implies the comprehension of the existence of false belief [18].

3. Appearance-Reality Distinction and False-Belief Understanding: What Types of Relationships?

Extensive research has identified an interesting period for ToM ability improvements between 3 and 5 years [14, 23]. Despite the recognized existence of significant individual differences [24, 25], cultural variations and task manipulations, and the earliest understanding of their own false belief than that of the other [26], a wide-ranging improvement in preschool age leads to support the idea that an important conceptual change occurs in this period [27]. Evidence in literature (see, [28, 29]) shows a similar pattern of development of false-belief and appearance-reality distinction understanding with significant correlations emerging between them. This finding is not surprising because both of them involve the same ability to recognize and cognitively manage conflicting representations [16]. Regarding false-belief understanding, most preschoolers fail at age 3 but subsequently grasp the correct answers at around age 5. However, the introduction of changes in the typical procedures of standard tests leads to increased performances of children under the age of 4; for

example, the children’s performance improves when they are actively involved [30, 31]. Both for FB tasks and AR tasks, 3- and 4-year-old children rarely distinguish between the way in which an object appears and the way in which it truly is [32]. Failing to pay attention to the two representations of the object (appearance and reality), they refer only to one of them. Therefore, a tendency to “phenomenism pattern error” is found when participants pay attention only to appearance. Instead, a tendency to “intellectual realism pattern error” occurs when they focus exclusively on the real characteristics of the object, regardless of how it looks [14]. The kind of thinking found in children of this age recalls the “irreversibility” described by [33], as part of the “preoperative stage.” At 3 years old children produce fewer correct answers and mostly realist errors; then at 4 years, errors range from 40% to 60%, mostly of the realist type [34]. Instead, the ability to recognize and distinguish appearance from reality seems mastered by 5-year-old children [16]. At this age, they quite clearly recall the reversibility of thought, and they are able to simultaneously perform more mental representations and integrate them as a whole, having in mind which are true and which are false. Findings in literature show that the period between the ages of 3 and 5 is significant, because children are constructing a new conceptual awareness that their mind and world are separate, and furthermore, the mind may misrepresent the true state of the world [15, 35]. In appearance-reality distinction the child has to be able to say “this looks like...but really is...” similar to the ability to understand other people’s false belief for which they should be able to say “he/she thinks this is...but I think this is...” As reported by Gopnik and Astington [16], interesting relationships between FB and AR understanding emerge: (1) both tasks require one to consider two conflicting representations; (2) both seem to develop between 3 and 5 years; and (3) both are correlated. Even if researchers find a similar pattern of development and significant relationships between false-belief and appearance-reality tasks, literature (see [29, 36]) also highlights developmental lags. In particular, two developmental patterns emerge [29]: (1) AR tasks are performed with more success than FB tasks (75% of children); (2) FB tasks are performed more easily than AR tasks (25% of children). These findings lead to questioning their meaning. One contribution in this direction is advanced by Melot and Angeard [37]. In their training study, they find a symmetrical transfer from FB to AR and from AR to FB that supports interdependence of the two constructs and isomorphy of the metarepresentational process they call into play. They trained preschool children in theory of mind tasks: two experimental groups were created: one trained in false-belief understanding and the other in the appearance-reality distinction. Children belonging to experimental groups were evaluated on appearance-reality and false-belief tasks and given explanations and feedback on their performance during two training sessions. Children belonging to the control group also were evaluated on false-belief and appearance-reality tests, but they received no feedback. At posttest, children in the control group showed no improvement suggesting that explanations and feedback are necessary for improvement. Instead, the two experimental groups showed a direct effect on the trained task (false-belief

or appearance-reality task) but also a transfer effect of the benefits of the training on the task that was not being trained (i.e., on the appearance-reality test in the false-belief group) highlighting the interdependency between these two concepts.

4. Rationale for This Study

Literature highlights the onset of theory of mind (ToM) in preschool time [38]. For this reason, we decided to conduct this study focusing on that interesting period. However, two important points should be underlined: firstly, even though this last aspect of the theory of mind construct has been widely investigated, AR distinction understanding remains in the background. Secondly, the debate is still open about the relationship between AR distinction and FB understanding. Our study moved from those two key points and aimed firstly to investigate how successes and failures change in 3-, 4-, and 5-year-old children's performances in AR task and secondly to investigate their relationships with performances on unexpected location, deceptive content, and deception comprehension tasks. In addition, we used a set of FB tasks that allow us to highlight differences in performance through the use of puppets and pictures (see "Sally and Ann" task) and differences in performance comparing one's own and others' false-belief understanding (see "Smarties" task) and how they relate to the understanding of AR distinction.

5. Research Questions and Hypotheses

This study aims to

- (1) evaluate the performance of 3-, 4-, and 5-year-old children in an appearance-reality task;
- (2) investigate relationships between results to the appearance-reality task (Task 1) and those obtained in the following tasks:
 - (i) unexpected location, "Sally and Ann" with puppets false-belief task (Task 2a);
 - (ii) unexpected location, "Sally and Ann" with pictures false-belief task (Task 2b);
 - (iii) deceptive contents, comprehension of our own false belief via the "Smarties task" (Task 3a);
 - (iv) deceptive contents, comprehension of others' false beliefs via the "Smarties task" (Task 3b);
 - (v) deception comprehension task (Task 4).

Regarding the first aim, we expected significantly low comprehension of the distinction between appearance and reality for 3-year-old children and better performance for 5-year-old children, with 4-year-old children improving significantly.

Regarding the second aim, in line with literature [29, 36] we expected an improvement between 3 and 5 years to cross all the tests, with a higher proportion of children that better comprehend AR distinction than FB. Furthermore, in line with literature [30, 31], we assumed that the use of puppets in false-belief task (task 2a) could facilitate a greater

understanding and better performance thanks to their capacity to stimulate interest in children. So, we predicted that performance on the "Sally Ann" with puppets false-belief task (task 2a) [25] would yield better performance than those obtained using the pictures (task 2b), especially for younger children, for whom the understanding of the situation in the pictures may be more complex than its staging using puppets.

6. Method

6.1. Participants. Ninety-one children (48 females and 43 males, age range = 3–5 years; $SD = .87$) were randomly selected from a preschool located in the suburbs of a large city of central Italy. Children were 3 ($N = 37$, M -age = 3.4 years), 4 ($N = 23$, M -age 4.5 years), or 5 years old ($N = 31$, M -age 5.4 years).

Children with certified disabilities were not included. Parents and school authorities, as well as the children themselves, gave consent to participate in the study. According to school officials, the socioeconomic level of the participants ranged from lower-middle class (85%) to upper-middle class (15%). The socioeconomic level was documented on the basis of their parents' qualifications and employment.

This research was endorsed by the Departmental Ethics Committee, Department of Education and Psychology, University of Florence. The authorities also gave consent to participate in the study.

6.2. Procedure, Measures, and Coding System. The following tasks were individually administered to the children in a quiet space in the school, but outside the classroom.

6.3. Task 1: Appearance-Reality (AR) Task [14]. To assess the participants' ability to distinguish the difference between appearance and reality, the appearance-reality task was used [14]. In a preliminary session prior to the examination with other children, we observed that children rapidly learn how to respond to this kind of test, so we prefer to use a single measure of AR.

In AR task the color of an object changes by means of a colored filter. In the test administered for this research a red glass containing milk was used. The glass of milk was placed on a table so that the participant could see the contents but could not look inside. The experimenter showed each subject the glass saying that it contained milk.

Then, (s)he posed the following question: "Tell me, what color is the milk in this glass?" The answer may be "the milk is white, but in this glass, it appears red" or something similar. If the participant answered the first question saying "red," the researcher asked the control question: "Truly, what color is the milk?"

If the participant answered the first question saying "white," the researcher asked the control question: "What color do you see the milk in this glass?"

All responses were recorded, transcribed, and then coded. Answers were coded as follows.

- (i) A score of 0 was assigned when the participant did not recognize the difference between appearance and

reality, that is, when they gave wrong answers to both questions “red-red” or “white-white.”

- (ii) A score of 1 was assigned when (a) the participant responded “red” to the first question and “white” to the second; (b) the participant responded “white” to the first question and “red” to the second, showing that, despite the perceptual salience of the appearance, the participant managed to keep in mind also the real representation; (c) the participant clearly retells his/her understanding of the difference between appearance and reality with expressions like “. . . it looks red because it is inside the red glass, but the milk is white.”

Interrater reliability was good (Cohen’s $k = .96$).

6.4. Task 2a: Task of False-Belief “Sally and Ann” Presented with Puppets [39]. For this false-belief task puppets were used to play the experimental situation of “Sally and Ann.” Characters were represented with two puppets named Maria and Francesco. The experimenter told the participant that Maria wanted to go for a walk and before leaving she put a ball in a colored box. Maria was then moved under the table so as to make her absence and inability to see what would happen clear. Another puppet, Francesco, takes the ball from the box and puts it in the trash. Finally, Maria returns and the participant is asked the false belief, “Where will Maria look for the ball?” The experimenter then asked the participant, “Well, why doesn’t Maria look in the . . . why wouldn’t Maria look there?” This additional question helped clarify the participant’s understanding of false belief. Responses were recorded, transcribed, and then coded.

Answers were coded as follows.

- (i) A score of 0 was assigned when the participant had not recognized the false belief.
- (ii) A score of 1 was assigned when the participant had demonstrated that they understood the false belief, either by giving the correct answer, or by giving an exact motivation despite a wrong answer.

Interrater reliability was good (Cohen’s $k = .98$).

6.5. Task 2b: Test of False-Belief “Sally and Ann” Presented with Pictures [39]. For this false-belief task pictures were used to deliver the “Sally and Ann” story. This picture variant helps determine whether this different presentation is more/less cognitively complex than the first puppet variant. Participants were presented with pictures and were read the accompanying text, which told a parallel story to the one previously presented. The test ended by asking the participant the false-belief question, “Where will Maria look for her ball?” as well as answering the motivation question, “Well, why doesn’t Maria look in the . . . why wouldn’t Maria look there?” Responses were recorded, transcribed, and then coded.

Answers were coded as follows.

- (i) A score of 0 was assigned when the participant had not recognized the false belief.

- (ii) A score of 1 was assigned when the participant had demonstrated that they understood the false belief, either by giving the correct answer or by giving an exact motivation despite a wrong answer.

Interrater reliability was good (Cohen’s $k = .98$).

6.6. Tasks 3a and 3b: Test of “Unexpected Content” or “Smarties Test” [4, 40]—Comprehension of Our Own and Others’ False Beliefs. The “Smarties test” was used to evaluate the participants’ ability to understand their own and others’ false beliefs. The task was administered using a commonplace tube of candy that actually contained pencils. The experimenter showed the closed tube of candies to the participant and questioned him/her about its contents. After the participant had said that the tube contained candies, the experimenter asked the participant to open the tube in order to check its contents. Once the participant discovered that the tube contained pencils, the investigator explained, “I finished the candies, so I used it for pencils,” and then two questions were asked, the first was about someone else’s false belief while the second was about their own false belief. The first question was “If we now call another child, we show them the box and ask them what’s in it, what would they say is inside the box?” The second question was “Before we opened the box, what did you think was in there?” Responses were recorded, transcribed, and then coded.

Answers were coded as follows.

- (i) A score of 0 was assigned when the participant had not recognized the false belief.
- (ii) A score of 1 was assigned when the participant demonstrated that they understood the false belief, either by giving the correct answer, or by giving an exact motivation despite a wrong answer.

Interrater reliability was good (Cohen’s $k = .96$).

6.7. Task 4: Deception Comprehension Task [8]. To understand each participant’s comprehension of deception, child listened to a story and then (s)he retold it. In the story a situation of deception is told that is an indicator of theory of mind. The stories are also used to evaluate the participants who were asked to retell them.

The story was as follows.

Animal Story

Once upon a time, in a wheat field, a little sparrow was greedily pecking ripe wheat grains. A cat, attracted by the rustle of the bird’s wings, came up silently to the bird and pow! In one moment with his paw he grabbed the little sparrow’s tail. The cat was about to eat the little sparrow when it said: “Hey, hey! A real gentleman never begins to eat if he is not clean!” The cat let the little sparrow go and cleaned his muzzle with his paws. Then the clever little sparrow fr. . . fr. . . flew away immediately. So the cat understood he had been tricked and in his heart he swore not to be tricked

TABLE 1: Story structure coding [20].

Level	Definition	Score
	No telling	0
First level: nonstory	Simple description of actions without characteristics of narrative style	1
Second level: sketch story	Characteristics of narrative style, such as introduction of the setting and the main characters, are present, but both problem and resolution are missing	2
Third level: incomplete story	Elementary narrative structure without a central event	3
Fourth level: essential story	Nonessential structural elements are missing	4
Fifth level: complete story	All eight elements are present; only the title is considered optional	5

TABLE 2: Children's language of mind coding system.

Category	Description	Examples
Physiological state (F)	Terms representing physiological states that might influence our behavior (such as hunger or thirst)	Being hungry, eating, drinking, being born, being ill
Perceptual state (P)	Terms describing what and how we perceive the world	Watching, listening, smelling, recognizing, feeling bad, felling hot/cold, noticing
Positive emotions state (EP)	Terms describing our positive feelings and emotions	Happy, pretty, nice, kiss, caressing, cuddle, hug, like, caring
Negative emotions state (EN)	Terms describing our negative feelings and emotions	Sad, angry, annoyed, ugly, scared, crying, screaming, getting bored, worrying, complaining
Willingness state (V)	Terms describing what we want to achieve and do	Willing, can, hoping, achieving, letting, trying, looking for, ordering
Cognitive state (C)	Terms representing what we cognitively think	Knowing, thinking, understanding, remembering, forgetting, clever, paying attention, true, false
Moral state (M)	Terms representing our moral perspective	Good, having to, reprimanding, promising, giving thanks, recommending, obeying
Sociorelational state (SR)	Terms describing the relationships between characters	Joking, helping, alone, becoming friends, abandoning, tricking

again. So, since that day, cats always clean their muzzle after their meal and not before!

All the stories retold by children were recorded and transcribed for analysis.

Answers were coded as follows.

To evaluate the deception comprehension, we referred to [8]; see also Table 2. The children's narratives were assessed by two independent judges, assigning a dichotomous score.

- (i) A score of 0 was assigned when there was no comprehension of deception: it was assigned when the sequence of historical events that lead to the situation of deception was not exposed properly and that is evident when the language of mind is not used properly or is used in a manner not relevant to deception.
- (ii) A score of 1 was assigned in the presence of the correct exposure sequence of events that make up the situation of deception accompanied by an appropriate mental language explaining the comprehension of

deception and including other references to false belief.

Interrater reliability was good (Cohen's $k = .95$).

6.8. Narrative Competence. Children's narrative competence strongly associated with ToM [41] was tested as a similar indicator to linguistic and verbal skills. Stories were evaluated in terms of structure, cohesion, and coherence in producing stories.

To analyze story structure, we used [42]. The presence, absence, and/or combinations of the eight fundamental elements (title, conventionalized story opening, characters, setting, problem, central event, resolution, and conventionalized story closing) allowed for rating of the stories into five categories, indicating varying levels of structural complexity, as shown in Table 1. Agreements between the judges (which was measured with Cohen's kappa, $k = .98$) were good. To analyze levels of cohesion in stories, the categories proposed by Halliday and Hasan [43] were used in order to detect cohesion among the elements of the story (e.g., the, thus, because, so, for, that, and consequently) and temporal cohesiveness,

TABLE 3: Statistical association between different comprehension of appearance-reality tasks and the age of the participants: Chi-square test.

Task	Comprehension	3 years	4 years	5 years	Chi-square
Appearance-reality ⁽¹⁾	Absent	18 (5.5)	1 (-1.8)	0 (-3.9)	29.56***
	Present	13 (-5.5)	13 (1.8)	27 (3.9)	
False-belief task with puppets	Absent	26 (3.8)	11 (-.5)	11 (-3.3)	15.90***
	Present	4 (-3.8)	9 (.5)	19 (3.3)	
False-belief task with pictures	Absent	22 (2.8)	13 (.7)	8 (-3.5)	13.41**
	Present	7 (-2.8)	8 (-.7)	20 (3.5)	
Smarties test 1	Absent	20 (3.3)	8 (-.2)	5 (-3.2)	13.20**
	Present	11 (-3.3)	12 (.2)	23 (3.2)	
Smarties test 2	Absent	23 (4.9)	6 (-1.1)	3 (-4.0)	25.83***
	Present	8 (-4.9)	14 (1.1)	25 (4.0)	

Note. Frequencies and residuals in bold are significant. Standardized adjusted residuals are in brackets. (1) Fisher Exact Test was used instead of Chi-square test. ** $p < .01$ and *** $p < .001$.

indicating a chronological sequence in the story (e.g., once upon a time, when, never, before, at the end, and suddenly). The amount of cohesiveness used by the participants, in proportion to the number of words produced, led to four increasing levels of cohesion: absent, low, medium, and high, corresponding to scores ranging from 0 to 3. Interrater agreement was good (Cohen's $k = .98$). Finally, to assess global story coherence, the sentences in the retold and transcribed stories were identified and their agreement was detected [44]. The amount of incoherence, proportional to the total number of sentences, produced four score categories (ranging from 0 to 3), indicating growing levels of coherence (absent, low, medium, and high). Interrater agreement reliability was good (Cohen's $k = .95$).

7. Data Analyses

Given the reduced variability of the scores, distribution frequencies were considered on a nominal dichotomic scale. In relation to the first aim we verified the developmental trend for the performances obtained on all the tasks through the construction of contingency tables which were created with two entries: "test" and "age." Chi-square tests were used to analyze differences and Fisher's Exact Test was used whenever the Chi-square was inappropriate. Furthermore, several Fisher's Exact Tests were performed to investigate the comprehension of the false belief in the different tasks for each age group. Several comparisons were made on the standardized adjusted residuals calculated in order to better understand the relationship between the variables. Regarding the second aim, adjusted standardized residuals of different contingency tables were carried out.

8. Results

To summarize the data on developmental trend and to introduce those on the comparison between tests, Table 1 reports the results obtained in each test for all three age groups considered.

Regarding the first aim, Table 1 results demonstrate a significant difference between appearance and reality [Fisher's Exact Test = 29.56, $p < .001$]. Standardized adjusted

residuals show a significant increase of comprehension for the appearance/reality task for 5-year-old children (presence: std. residual = 3.9) than for 3-year-old children (present: std. residual = -5.5) (Table 3).

Results demonstrate significant differences between the age groups; in fact, "false-belief task with puppets" [$\chi^2(2) = 15.90, p < .001$], "false-belief task with pictures" [$\chi^2(2) = 13.41, p < .01$], "Smarties test 1" [$\chi^2(2) = 13.20, p < .01$], and "Smarties test 2" [$\chi^2(2) = 25.83, p < .001$] show significantly different distributions of their scores in the three ages. Regarding all the false-belief tasks, the standardized adjusted residuals show a constant increase of the comprehension of the false belief as the age increases, with lack of comprehension being more at 3 years and a more frequent presence of comprehension at 5 years (Table 3). Unlike the appearance/reality task, the analysis of the standardized residuals shows that the difference in the comprehension of the false-belief tasks is between the 3- and the 5-year old children.

Concerning the comparison between the comprehension of the appearance-reality concept in the different tasks separately for each class of age (3 years old vs. 4 years old vs. 5 years old), for the 3-year-old children, comprehension performance of the tasks did not differ [Fisher's Exact Test = 7.28, $p = n.s.$] (Table 4), while for 4- and 5-year-old children, the comprehension of the appearance-reality task was more present than the comprehension for all the others tasks, respectively [Fisher's Exact Test = 13.46, $p < .01$], and [Fisher's Exact Test = 16.45, $p < .01$] in particular for false-belief task with puppets (Table 4). No significant difference was found for the deception comprehension task for the 3 age groups [Fisher Exact Test = 1.10, $p = n.s.$] (see Table 5).

Regarding the second aim, the comparison between "appearance-reality task" and all the false-belief tasks showed a significant positive association between "appearance-reality task" and all the false-belief tasks; that is performance significantly improved with participant age. The unique difference pointed out by the statistical analysis was that, regarding appearance/reality task, the increase of the comprehension was already localized between the 3- and 4-year-old children, whereas comprehension/performance for the false-belief task increased between 3- and 5-year-old children.

TABLE 4: Statistical association for appearance-reality and other tasks administered, by participant: Chi-square test.

Age	Comprehension	Appearance-reality	False-belief task with puppets	False-belief task with pictures	Smarties test 1	Smarties test 2	Fisher's Exact Test
3 years	Absent	18 (-1.9)	26 (2.0)	22 (.6)	20 (-1.0)	23 (.3)	7.28
	Present	13 (1.9)	4 (-2.0)	7 (-.6)	11 (1.0)	8 (-.3)	
4 years	Absent	1 (-2.8)	11 (1.4)	13 (2.2)	8 (-.1)	6 (-1.1)	13.46**
	Present	13 (2.8)	9 (-1.4)	8 (-2.2)	12 (.1)	14 (1.1)	
5 years	Absent	0 (-2.8)	11 (2.7)	8 (1.4)	5 (-.2)	3 (-1.3)	16.45**
	Present	27 (2.8)	19 (-2.7)	20 (-1.4)	23 (.2)	25 (-1.3)	

Note. The frequencies and the residuals reported in bold type are significant. Standardized adjusted residuals are bracketed. ** $p < .01$ and *** $p < .001$.

TABLE 5: Statistical association between “deception comprehension task” and participant: Fisher's Exact Test.

Task	Comprehension	3 years	4 years	5 years	Fisher's Exact Test
Deception comprehension task	Absent	6 (1.0)	11 (.5)	18 (-1.2)	1.10
	Present	0 (-1.0)	1 (-.5)	4 (1.2)	

Note. Standardized adjusted residuals are reported in brackets.

9. Discussion

This study aimed to highlight the different levels of performance in 3-, 4-, and 5-year-old children in an appearance-reality task. Findings revealed a first key aspect that, beyond a significant increase in performance at the appearance-reality task, 4-year-old children's performances reflect a still implicit awareness, as none of them, and even among children of five years, explicitly said “the milk seems red but actually it is white,” confirming the findings of other studies [14, 30, 34]. Our first hypothesis was confirmed, since children's answers are distributed in a significantly different way in relation to age: there is a tendency for which the wrong answers, particularly high at 3 years, greatly decrease towards 4 years. At 5 years, there is a tendency for answers that indicate the presence of appearance-reality distinction understanding. These results support the hypothesis that in preschool children a substantial improvement in performance is found between 3 and 5 years. 3-year-old children failing in appearance-reality distinction show a similar pattern of answers defined by [14] as “phenomenism error” and “intellectual realism error.” In the first case, children answer both questions only in terms of appearance “the milk in the glass appears red and is red.” Instead, in the second case, they answered only in terms of reality “milk is white and appears white.” These arrangements bring out the difficulty of a younger child to pay attention to both aspects simultaneously. Specifically, in our sample, the majority of 3-year-old participants produce “phenomenism” bias. For this purpose, it is important to consider that the test used, focusing on the characteristic of an object (the color), promotes the tendency to pay attention to the perceptually most salient aspect, appearance [45]. Only three children produced “intellectual realism” bias: they answered both questions in terms of reality, a trend that generally prevails in evidence concerning the identity, and not the perceptual ownership such as in this case, of an object [46]. Briefly, the results lead us to hypothesize that younger children have difficulty in mentally manipulating two conflicting

representations and they try to resolve this by focusing only on the most salient aspect. Regarding the sample of 5-year-old children, despite all of them having answered both questions correctly, succeeding in distinguishing appearance (red) from reality (white), none of them produced terms like “seems” or “looks like” to clarify their understanding. The answers that emerged to the question on appearance are as follows: our subjects respond to the question on appearance “it is red,” and never “seems red,” as well as the next question on reality, “it is white.” The expression “seems” is rarely indeed used by children of this age; however, there was a spontaneous use of words like “actually” and “truly” [47]. So, we could summarize by saying that 5-year-old children, although doing well in the tests, still tend to analyze objects and situations sequentially, considering appearance as a reality in a given moment. It is clear that these children have not yet fully acquired the reversibility of thought [33], but they have the basic skills to develop a more structured and conscious ability to distinguish and simultaneously manage all possible representations of an object or an event, with all the implications that this entails. The ability to consider and manage different representations, on the other hand, concerns not only recognition of the distinction between appearance and reality, but also the recognition of false belief. So, they show their understanding, but at the same time they are not able to fully master and simultaneously manage both conflicting representations. Referring to Wellman and Liu [48] we might think that children have an implicit understanding of the difference between appearance and reality, but they are unable to explain their understanding with the expression: “the milk in the glass seems red, but it is white.” However, the hypothesis that the difficulty of younger children is due to cognitive immaturity in the management of the double representation would be called into question because, observing spontaneous behavior in different familiar contexts, children show that they act as if they have real-apparent level in mind regarding the characteristics of the object or situation they encounter [49].

TABLE 6: *M* and *ds* of the deception comprehension task in narrative story and level of textual competence.

	3 years <i>M</i> (<i>ds</i>)	4 years <i>M</i> (<i>ds</i>)	5 years <i>M</i> (<i>ds</i>)
Structure	1.3 (0.5)	2.3 (1)	2.8 (1)
Coherence	0.2 (0.4)	1 (1)	1.4 (0.9)
Cohesion	1 (0)	1.1 (0.3)	1.3 (0.6)

The second purpose of the study was to investigate the relationship between AR performances and FB performances (unexpected location, deceptive content, and deception comprehension tasks performances). Results partially confirmed our hypotheses. An improvement between 3 and 5 years across all the tests emerged, but only children at 4 and 5 years show that they master AR distinction understanding better than FB comprehension in the other tests. These results seem in line with another finding (see [29, 36]) that highlights developmental lags between FB and AR task. In particular, our results overlap with the developmental patterns expressed by the majority of children (75%), according to which AR tasks are performed with more success than FB tasks. However, we must consider that the result informs us about an early onset of appearance and reality distinction understanding; in fact, no 5-year-old child spontaneously says “the milk in the glass seems red, but it is white.” If we had considered this answer, no one had a good test result. Regarding 3-year-old children, comparison between AR tasks and FB tasks shows a similar pattern of answers, in line with [28, 29]: those who correctly respond to the AR task also do so in FB task, and, on the contrary, those who fail tend to fail even in the others. In contrast to our assumptions, the results obtained with the two forms of FB task “Sally and Ann” [39] with puppets and with pictures, while showing a clear progression with age, do not show any performance difference between the two forms. The introduction of more familiar and interesting materials does not facilitate performance of younger children on the test, leading to hypotheses that the task is beyond the reach of the general cognitive level of young children.

Finally, we have to underline that nearly half of the participants did not want to tell the story after hearing it from the researcher. Most were children of 3 years. On the other hand, although the majority of 5-year-old children were able to tell the story proposed by the investigator, very few understood the deception in the narrative. 3-year-old children produced, on average, “nonstories,” while 4- and 5-year-old children produced, on average, “sketch stories” (see Tables 1 and 6). This progression is in line with literature where an improvement in the ability to tell and retell stories between 3 and 5 years is highlighted [50, 51]. Beyond narrative competence, results showed no improvement in performance in the deception understanding task nor from 3 to 5 years old. This finding leads us to suppose that all children have, likewise, encountered significant difficulty with this kind of task. We might think that the difficulty of one test added to that of the other, making comprehension of the deception task too hard even for 5-year-old children. According to

Siegal [52], in addition, some elements of the procedures used in standard tests, such as the type of questions, the kind of objects used, and the method of the tasks, might contribute to the difficulties shown in their conduct by younger children [47]. On the other hand, they also suggest that the difficulties arising for younger children may be due to the use, in all standard tests of verbal response mode. To explain these reactions, it may be useful to refer to the distinction between implicit and explicit understanding of theory of mind proposed by [53]. In line with this, the child senses that the other has a false belief and proves it implicitly; however, he/she shows difficulty in explaining this intuition, giving the wrong answer to the question on the false belief.

The importance of theory of mind is central in child development for its relevance to comprehension of the surrounding world. In particular, understanding of the difference between appearance and reality represents an important acquisition and so a useful role for the child’s future learning. The results obtained in this work might suggest that, in this transition period, ranging from 3 to 5 years, the school may operate in the “proximal development zone” [54] for children. One of the fields that lends itself to stimulate the understanding of appearance and reality is scientific learning, such as in subjects like biology and chemistry [55]. With this study, it was shown how 3-year-old children have difficulty in mastering two conflicting representations. The proposal to introduce science topics at school for children aiming to promote the development of a scientific “thinking” cannot be achieved by showing such small children only picture books containing experiments, natural phenomena, and expecting to teach scientific concepts such as chemistry or physics. At this age in fact children have not yet mastered a decentralized and reversible thought, necessary for the formation of scientific concepts and the construction and reconstruction of knowledge about the world, both physical and social [56]. Science learning in kindergarten can also have moments of observation of natural situations, where children spontaneously grasp the understanding of the difference between appearance and reality.

Future research could consider systematic observations on a larger sample in natural situations, as close as possible to a real-life situation or practice detections in spontaneous life conversations and in particular in natural situations of exploration of nature. A further limitation is that this study used only a single measure for AR distinction understanding, so future research should submit different groups of children to different AR tasks or they could set up a task that is impermeable to children’s learning of the answers to the tests.

Competing Interests

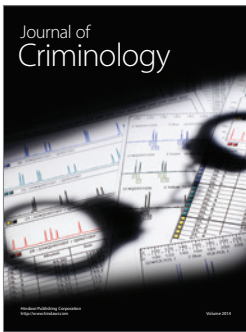
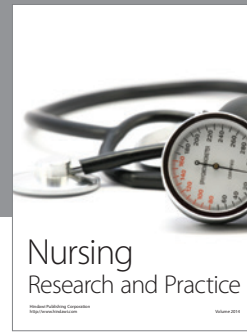
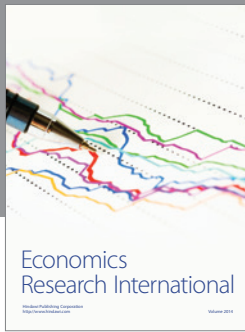
The authors declare that there is no conflict of interests regarding the publication of this paper.

References

- [1] D. Premack and G. Woodruff, “Does the chimpanzee have a theory of mind?” *Behavioral and Brain Sciences*, vol. 1, no. 4, pp. 515–526, 1978.

- [2] J. W. Astington, "Sometimes necessary, never sufficient: false belief understanding and social competence," in *Individual Differences in Theory of Mind: Implications for Typical and Atypical Development*, B. Repacholi and V. Slaughter, Eds., pp. 13–38, Psychology Press, New York, NY, US, 2003.
- [3] C. Duvala, P. Piolino, A. Bejanina, F. Eustachea, and B. Desgranges, "Age effects on different components of theory of mind," *Consciousness and Cognition*, vol. 20, no. 3, pp. 627–642, 2011.
- [4] Z. Wang and L. Wang, "The mind and heart of the social child: developing the empathy and theory of mind scale," *Child Development Research*, vol. 2015, Article ID 171304, 8 pages, 2015.
- [5] C. Hughes, S. Lecce, and C. Wilson, "Do you know what I want? Preschoolers' talk about desires, thoughts and feelings in their conversations with sibs and friends," *Cognition and Emotion*, vol. 21, no. 2, pp. 330–350, 2007.
- [6] H. M. Wellman, D. Cross, and J. Watson, "Meta-analysis of theory-of-mind development: the truth about false belief," *Child Development*, vol. 72, no. 3, pp. 655–684, 2001.
- [7] J. Perner and B. Lang, "Development of theory of mind and executive control," *Trends in Cognitive Sciences*, vol. 3, no. 9, pp. 337–344, 1999.
- [8] B. A. Gamannossi and G. Pinto, "Theory of mind and language of mind in narratives: developmental trends from kindergarten to primary school," *First Language*, vol. 34, no. 3, pp. 262–272, 2014.
- [9] K. Bartsch and H. M. Wellman, *Children Talk about the Mind*, Oxford University Press, New York, NY, USA, 1995.
- [10] G. Mazzoni, F. Chiesi, L. Bigozzi, and C. Alocci, "Parents' reports about children's early ability to lie," 1995.
- [11] J. W. Astington and J. Pelletier, "The language of mind: its role in teaching and learning," in *The Handbook of Education and Human Development*, D. R. Olson and N. Torrance, Eds., pp. 593–619, Blackwell, Oxford, UK, 1996.
- [12] Z. He, M. Bolz, and R. Baillargeon, "False-belief understanding in 2.5-year-olds: evidence from violation-of-expectation change-of-location and unexpected-contents tasks," *Developmental Science*, vol. 14, no. 2, pp. 292–305, 2011.
- [13] Z. He, M. Bolz, and R. Baillargeon, "2.5-Year-olds succeed at a verbal anticipatory-looking false-belief task," *British Journal of Developmental Psychology*, vol. 30, no. 1, pp. 14–29, 2012.
- [14] J. H. Flavell, E. R. Flavell, and F. L. Green, "Development of the appearance-reality distinction," *Cognitive Psychology*, vol. 15, no. 1, pp. 95–120, 1983.
- [15] B. Sodian, C. Taylor, P. L. Harris, and J. Perner, "Early deception and the child's theory of mind: false trails and genuine markers," *Child Development*, vol. 62, no. 3, pp. 468–483, 1991.
- [16] A. Gopnik and J. W. Astington, "Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction," *Child Development*, vol. 59, no. 1, pp. 26–37, 1988.
- [17] C. A. Miller, "False belief and sentence complement performance in children with specific language impairment," *International Journal of Language and Communication Disorders*, vol. 39, no. 2, pp. 191–213, 2004.
- [18] H. Wimmer and J. Perner, "Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception," *Cognition*, vol. 13, no. 1, pp. 103–128, 1983.
- [19] J. Perner, T. Ruffman, and S. R. Leekam, "Theory of mind is contagious: you catch it from your sibs," *Child Development*, vol. 65, no. 4, pp. 1228–1238, 1994.
- [20] A. G. Spinillo and G. Pinto, "Children's narratives under different conditions: a comparative study," *British Journal of Developmental Psychology*, vol. 12, no. 2, pp. 177–193, 1994.
- [21] T. L. Hutchins, P. A. Prelock, H. Morris, J. Benner, T. Lavigne, and B. Hoza, "Explicit vs. applied theory of mind competence: a comparison of typically developing males, males with ASD, and males with ADHD," *Research in Autism Spectrum Disorders*, vol. 21, pp. 94–108, 2016.
- [22] N. Epley, C. K. Morewedge, and B. Keysar, "Perspective taking in children and adults: equivalent egocentrism but differential correction," *Journal of Experimental Social Psychology*, vol. 40, no. 6, pp. 760–768, 2004.
- [23] S. Lecce, S. Zocchi, A. Pagnin, P. Palladino, and M. Taumoepeau, "Reading minds: the relation between children's mental state knowledge and their metaknowledge about reading," *Child Development*, vol. 81, no. 6, pp. 1876–1893, 2010.
- [24] A. L. Cutting and J. Dunn, "Theory of mind, emotion understanding, language, and family background: individual differences and interrelations," *Child Development*, vol. 70, no. 4, pp. 853–865, 1999.
- [25] C. Hughes and R. Ensor, "Executive function and theory of mind in 2 year olds: a family affair?" *Developmental Neuropsychology*, vol. 28, no. 2, pp. 645–668, 2005.
- [26] S. Brandt, D. Buttelmann, E. Lieven, and M. Tomasello, "Children's understanding of first- and third-person perspectives in complement clauses and false-belief tasks," *Journal of Experimental Child Psychology*, vol. 151, pp. 131–143, 2016.
- [27] H. M. Wellman, F. Fang, D. Liu, L. Zhu, and G. Liu, "Scaling of theory-of-mind understandings in Chinese children," *Psychological Science*, vol. 17, no. 12, pp. 1075–1081, 2006.
- [28] J. H. Flavell, F. L. Green, and E. R. Flavell, "Children's understanding of the stream of consciousness," *Child development*, vol. 64, no. 2, pp. 387–398, 1993.
- [29] A.-M. Melot, O. Houdé, S. Courtel, and L. Soenen, "False beliefs attribution, distinction between appearance and reality and visual perspective taking: do they develop simultaneously," *The Genetic Epistemologist*, vol. 23, 1995.
- [30] M. W. Battacchi, G. Celani, and A. Bertocchi, "The influence of personal involvement on the performance in a false belief task: a structural analysis," *International Journal of Behavioral Development*, vol. 21, no. 2, pp. 313–329, 1997.
- [31] C. Lewis, N. H. Freeman, C. Hagestadt, and H. Douglas, "Narrative access and production in preschoolers' false belief reasoning," *Cognitive Development*, vol. 9, no. 4, pp. 397–424, 1994.
- [32] G. O. Deák, S. D. Ray, and K. Brennenman, "Children's perseverative appearance-reality errors are related to emerging language skills," *Child Development*, vol. 74, no. 3, pp. 944–964, 2003.
- [33] J. Piaget, *Le Jugement et le Raisonnement chez l'enfant*, Delachaux et Niestlé, Neuchâtel, Switzerland, 1924.
- [34] J. H. Flavell, F. L. Green, K. E. Wahl, and E. R. Flavell, "The effects of question clarification and memory aids on young children's performance on appearance-reality tasks," *Cognitive Development*, vol. 2, no. 2, pp. 127–144, 1987.
- [35] H. M. Wellman, F. Fang, and C. C. Peterson, "Sequential progressions in a theory-of-mind scale: longitudinal perspectives," *Child Development*, vol. 82, no. 3, pp. 780–792, 2011.

- [36] A.-M. Melot and O. Komano, "Changes in theory of mind: a longitudinal study," *The Genetic Epistemologist*, vol. 25, 1997.
- [37] A.-M. Melot and N. Angeard, "Theory of mind: is training contagious?" *Developmental Science*, vol. 6, no. 2, pp. 178–184, 2003.
- [38] V. Slaughter and D. Perez-Zapata, "Cultural variations in the development of mind reading," *Child Development Perspectives*, vol. 8, no. 4, pp. 237–241, 2014.
- [39] U. Frith, *Autism: Explaining the Enigma*, Basil Blackwell, Oxford, UK, 1989.
- [40] A. Gopnik and V. Slaughter, "Young children's understanding of changes in their mental states," *Child Development*, vol. 62, no. 1, pp. 98–110, 1991.
- [41] G. Pinto, C. Tarchi, B. Accorti Gamannossi, and L. Bigozzi, "Mental state talk in children's face-to-face and telephone narratives," *Journal of Applied Developmental Psychology*, vol. 44, pp. 21–27, 2016.
- [42] G. Pinto, L. Bigozzi, B. A. Gamannossi, and C. Vezzani, "Emergent literacy and learning to write: a predictive model for Italian language," *European Journal of Psychology of Education*, vol. 24, no. 1, pp. 61–78, 2009.
- [43] M. A. K. Halliday and R. Hasan, *Cohesion in English*, Longman, London, UK, 1976.
- [44] L. R. Shapiro and J. A. Hudson, "Tell me a make-believe story: coherence and cohesion in young children's picture-elicited narratives," *Developmental Psychology*, vol. 27, no. 6, pp. 960–974, 1991.
- [45] C. Heyes, "False belief in infancy: a fresh look," *Developmental Science*, vol. 17, no. 5, pp. 647–659, 2014.
- [46] J. H. Flavell, "The development of children's knowledge about the appearance-reality distinction," *American Psychologist*, vol. 41, no. 4, pp. 418–425, 1986.
- [47] J. H. Flavell, E. R. Flavell, and F. L. Green, "Development of the appearance-reality distinction," *Cognitive Psychology*, vol. 15, no. 1, pp. 95–120, 1983.
- [48] H. M. Wellman and D. Liu, "Scaling of theory-of-mind tasks," *Child Development*, vol. 75, no. 2, pp. 523–541, 2004.
- [49] F. Sapp, K. Lee, and D. Muir, "Three-year-olds' difficulty with the appearance—reality distinction: is it real or is it apparent?" *Developmental Psychology*, vol. 36, no. 5, pp. 547–560, 2000.
- [50] G. Pinto, C. Tarchi, and L. Bigozzi, "Development in narrative competences from oral to written stories in five- to seven-year-old children," *Early Childhood Research Quarterly*, vol. 36, pp. 1–10, 2016.
- [51] L. Bigozzi and G. Vettori, "To tell a story, to write it: developmental patterns of narrative skills from preschool to first grade," *European Journal of Psychology of Education*, vol. 31, no. 4, pp. 461–477, 2016.
- [52] M. Siegal, *Knowing Children: Experiments in Conversation and Cognition*, Erlbaum, Hove, UK, 1991.
- [53] W. A. Clements and J. Perner, "Implicit understanding of belief," *Cognitive Development*, vol. 9, no. 4, pp. 377–395, 1994.
- [54] L. S. Vygotsky, *Mind in Society: The Development of Higher Psychological Processes*, Harvard University Press, Cambridge, Mass, USA, 1980.
- [55] L. Bigozzi, C. Vezzani, C. Tarchi, and C. Fiorentini, "The role of individual writing in fostering scientific conceptualization," *European Journal of Psychology of Education*, vol. 26, no. 1, pp. 45–59, 2011.
- [56] L. Bigozzi, A. Biggeri, F. Boschi, P. Conti, and C. Fiorentini, "Children 'scientists' know the reasons why and they are 'poets' too. Non-randomized controlled trial to evaluate the effectiveness of a strategy aimed at improving the learning of scientific concepts," *European Journal of Psychology of Education*, vol. 17, no. 4, pp. 343–362, 2002.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

