Cognitive Development 59 (2021) 101091

Contents lists available at ScienceDirect

Cognitive Development

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

Age 4 transitions: Reflection as a domain-general development for explicit reasoning

Jedediah W.P. Allen[a,*], Bartuğ Çelik[a], Mark H. Bickhard[b]

[a] Department of Psychology, Bilkent University, Ankara, 06800, Turkey
[b] Departments of Philosophy & Psychology, Lehigh University, Bethlehem, PA, 18015, USA

ARTICLE INFO

Keywords:
Interactivism
Action-based approaches
Reflection
Mental imagery
Domain-general development
Preschoolers

ABSTRACT

The literature provides many examples of important developments across different social and cognitive domains at around age 4. Based on an action-based approach to cognition - interactivism - we argue that the changes across different domains can be explained by the development of a domain-general cognitive enabling: *reflection*. The interactivist model's claim about reflection was empirically supported on the basis of a novel object-reasoning task called Leaning Blocks (LB) developed by Allen and Bickhard (2018). In the current study, there were three aims. First, to replicate the age 4 shift on the LB task in a non-western sample. Second, to explore the LB task's relations with Executive Functioning skills (working memory, cognitive flexibility, and inhibitory control) in order to test alternative interpretations of performance on LB. Third, to diversify the measurement of reflective reasoning with two new tasks based on Piaget's categorization of mental imagery. The results showed a replication of Allen and Bickhard's 2018 findings on the LB task and failed to show any relations with the EF measures after controlling for age. One of the new reflection tasks (Candy Monster) showed both the age 4 transition and a correlation with LB (before-and-after controlling for age and the EF measures). Overall, as in Allen and Bickhard (2018) the current study's results support the interactivist model's claim that the development of reflection is responsible for the transitions in performance across domains in a "stage-like" fashion.

1. Introduction

Age 4 transitions in performance seem to exist for a variety of tasks within and across different developmental domains. Empirical results have demonstrated important developments at around age 4 for abilities in a diversity of areas that include false-belief understanding (Wellman, Cross, & Watson, 2001); appearance-reality reasoning (Flavell, Flavell, & Green, 1983); aspects of language (de Villiers & Pyers, 2002); learning from others (Harris, 2012); deception (Lee, 2013); autobiographical memory (Nelson & Fivush, 2004); and counterfactual and future thinking (Atance, 2008; Guajardo, Parker, & Turley-Ames, 2009) as well as for domain-general abilities such as Executive Functions (EF, Zelazo, Müller, Frye, & Marcovitch, 2003); meta-cognition (Woolley & Ghossainy, 2013) and the self-regulation related to delay behavior (Mischel, Shoda, & Rodriguez, 1989). The relative synchrony of such changes in performances suggests the possibility of a domain-general development as an explanation.

In an earlier study, Allen and Bickhard (2018) have argued that epistemic reflection constitutes the domain-general development

* Corresponding author.
E-mail address: jallen@bilkent.edu.tr (J.W.P. Allen).

https://doi.org/10.1016/j.cogdev.2021.101091
Received 21 October 2020; Received in revised form 7 July 2021; Accepted 15 July 2021
Available online 29 July 2021

that is responsible for explaining the relative synchrony in performances on the variety of tasks. To support their conclusion, the authors developed a novel object-reasoning task called Leaning Blocks (LB). Children showed an age 4 transition on the new task such that 3-year-olds were different than 4- and 5-year-olds but the later two groups were not different from each other. Allen and Bickhard argued that the threshold effect in performance at age 4 was due to the minimal performance factors needed to pass the task (i.e., that it was a relatively "pure" measure of reflection). However, measures of potential performance factors were not included and tested statistically. In particular, children's performance on the LB task may involve the well-documented developments in EF skills (i.e., cognitive flexibility, working memory, inhibitory control). Accordingly, in the current study we sought to replicate and extend the original findings from Allen and Bickhard as well as to test for other domain-general processes (i.e., EF skills) that are known to undergo important developments at around age 4.

From the current perspective, we argue that age 4 transitions are accounted for by newly developed reflection abilities that are domain-general and enable the development of qualitatively new forms of knowing. Despite some points of convergence with those arguing for domain-general developments around age 4 (Halford, Wilson, & Phillips, 1998; Perner, 1991; Zelazo, 2004), the current approach (i.e., interactivism, see Bickhard, 2009a, 2009b, in-preparation) is situated squarely within an action-based constructivist tradition that includes Piaget as one of its best known practitioners. Further, from the interactivist perspective (and in contrast to information-processing approaches), the dynamics of knowledge construction at age four cannot be separated from the domain-general processes (reflection) that enable the qualitatively new forms of knowing to develop.

1.1. Interactivism

Action-based approaches argue that the foundations of cognition and mind are constituted by (inter)action with the physical and social environment. From such a perspective, infants do not possess representations and then use them to act on the world as the output of computations. Instead, interaction potentialities *are* the representations (see below for elaboration). Interactivism is a well-developed action-based approach to cognition and mind that provides multiple specific models for representation and knowing, perception, language, emotions, and other cognitive and social phenomena (Bickhard, 2009a, b, in-preparation; Bickhard & Terveen, 1995). For the interactivist framework, anticipation of potential interactions forms the core of the model of "mental" representation (Bickhard, 1980). That is, animals and infants know the world in terms of their anticipations for possible (inter)action. In turn, appropriate differentiation must take place for those anticipations to have the possibility of success. For example, if a frog in the lab correctly/reliably anticipates an opportunity to flick its tongue and eat, then it will need to be able to differentiate flies from pebbles. Similarly, an infant must learn to differentiate its physical and social environments in order to learn the very different sorts of anticipations related to each.

From this perspective, development is an on-going process of organizing internal interconnected webs of anticipations in ways that allow for important abilities to develop. For example, Piaget's model of object representation as assimilated into the modeling resources of the interactivist approach (such assimilation is possible due to the shared pragmatist foundation of both genetic epistemology and interactivism, Bickhard & Campbell, 1989). By 24 months of age, representation of small manipulable objects is constituted by an organized web of anticipations for potential interaction. This organization has two important properties. First, any point in the web is reachable from any other point such that anticipations about visual scans, for example, are reachable via anticipations about rotating the object with one's hands. Second, the organization of anticipations remains invariant under a large class of other manipulations. For example, if the object is occluded or displaced, the organization of anticipations remains available through relevant intervening steps. Such invariance, however, is not absolute. If the object is crushed or burned, then the web of potential interactions no longer exists.

Object permanence then, is an emergent property of the internal organization of the toddler's anticipations. More broadly, the development of object permanence is itself a development of the toddler's differentiation of themselves from the world – an understanding that the world exists independent of their actions. This understanding allows for a sort of distancing between "thought" and action but it is a *functional* (and not an *epistemological*) distancing. That is, the distinction between self and world is an implicit differentiation in the toddler's way of knowing but the distinction itself is not known by the toddler. An explicit understanding of the self-world differentiation awaits the development of reflection. Reflection can be understood as a type of abstraction process that does not depend on language. For interactivism, reflection must develop to enable moving beyond strictly interactive knowing and such development must involve an architectural change to the organism (Bickhard, 1980). In humans, this architectural change is the end point of a macro-evolutionary ratchet that builds on a sequence from interactive knowing to learning to emotions and then to reflection. Accordingly, the development of reflection is thought to involve changes in the brain[1] that are largely under maturational control. Age 4 has been established as the period for such maturation on empirical grounds (Bickhard, 1992).

From the interactivist perspective, prior to age 4, all knowing is through interaction with the world (Bickhard, 1980, 1992; Campbell & Bickhard, 1986). Properties and relations that are not concretely available for interaction may exist implicitly in the organization of anticipations related to the concrete interactions. Only with the possibility of reflection, may children come to know explicitly about such abstract properties and relations. One consequence of this model of representation and reflection is that objects are not known explicitly until age 4. Although contrary to most other perspectives, this means that infants and toddlers do not yet

[1] While not focused on the developmental aspects per se, Bickhard (2015b) discusses how reflection might operate at the level of the brain in terms of internal (oscillatory) interaction-supporting (looping) connections from, e.g., prefrontal to basal ganglia to thalamus to prefrontal. Likely 'late' maturing would be prefrontal per se and basal ganglia to thalamus.

represent objects qua objects. Making sense of this claim seems to require an action-based approach (e.g., would seem to require a model of implicit representational content, see Bickhard, 1998). In contrast, if the nature of representation is exhausted by the theoretical resources of encoding relationships (e.g., symbols or activation patterns as constituting informational, causal, or isomorphic relationships), then implicit content is not available to be reflected upon (Bickhard & Terveen, 1995).

1.1.1. Leaning blocks (LB)

The LB task was designed to assess reflection in an area that is not usually assumed to be undergoing new developments during preschool (i.e., the object concept/representation).[2] The task seeks to differentiate between children who only possess implicit presuppositions about object properties and relations from children who can explicitly represent such presuppositions. The difficulty for designing such a task is that any particular procedure (i.e., behavioral assessment) can be accomplished through interactive knowing alone (i.e., without reflection). This means that the task must involve sufficient novelty such that the child has not learned pre-reflective procedures for passing it. From the current perspective, pre-reflective children are assumed to represent objects in terms of an organization of anticipations that remains invariant under a large class of other transformations. This organization is not known by the pre-reflective child; instead, pre-reflective children know how to interact with objects in ways that presuppose features of the organization (e.g., the relative permanence of objects).

In the LB task, children are asked to answer a question which requires them to reason about the relation between two objects – specifically, the mutually supportive relationship of two blocks leaning against each other. The overall task has three steps and three trials. In the first step, children are asked whether a red block that is held on a 45 degree angle from the floor will "fall" or "stay up". After children respond, the block is released. Then a second red block is held opposite the location of the first block and children are asked the same question after which the block is released. Finally, the two blocks are held opposite of each other such that they are mutually supportive. Before removing one's hands, children are again asked whether the blocks will fall or stay up. It is this third question that we argue *requires* reflection (i.e., in order to reason about the novel relationship between the two blocks). This 3-step sequence is repeated two more times using other materials (i.e., two large rectangular pieces of cardboard and two smaller white blocks). Results from Allen and Bickhard (2018) indicated that 3-year-olds could generally answer the first two questions correctly but not the third for all three trials; while 4- and 5-year-olds could answer all three questions correctly. Further, there was evidence to suggest that this age 4 transition on the LB task was the consequence of a domain-general development. Specifically, performance on LB was correlated with three different false-belief tasks (change of location, unexpected contents, and active deception).

1.2. Types of image

The LB task can be characterized in terms of Piaget's classification of different types of images which includes the distinction between reproductive and anticipatory imagery. Reproductive images concern situations in which the child has had previous experience with the content of the image (e.g., falling objects). Anticipatory images concern situations in which the child has not experienced the content before (e.g., mutually supported objects). Reproductive and anticipatory images can be further separated into two categories, "kinetic" and "transformational". Kinetic images are defined in terms of whether the images include an object's displacement without any change in its form (e.g., moving a block). In contrast, transformational images are defined to include an object's movement while also changing its form (e.g., smashing a block; see Piaget & Inhelder, 1971 for more detail).

The LB task is an *anticipatory-kinetic* imagery task in which participants need to reason about the displacement of the object(s) without any change to the object(s) itself. The current study also developed a simplified version of an *anticipatory-transformational* imagery task that was originally created by Piaget and Inhelder (1971, pp. 215-218). The Elastic Band (EB) task requires participants to reason about changes to the shape of an elastic band that is stretched over a configuration of wooden pegs (e.g., a square, involving four pegs, changes to a triangle or straight line with the removal of one and two pegs respectively). The EB task was included in the current study because it provided for an opportunity to compare children's performance on the two types of anticipatory imagery (*kinetic* for LB versus *transformational* for EB). A new reflection task (Candy Monster, CM) was also developed that is intended to be another *anticipatory-kinetic* imagery task similar to the LB task.

1.3. Alternative domain-general explanations of leaning blocks and false-belief tasks

Although the claims about epistemic reflection and its taking place around age 4 go back some years (Bickhard, 1980, 1992; Campbell & Bickhard, 1986); direct empirical evidence for it is relatively new (Allen & Bickhard, 2018). Further, there are well established developments in EF abilities around age 4 that could also be thought to explain many of the age 4 transitions mentioned above. Accordingly, it is worthwhile to test whether these EF abilities might account for the age 4 transition found for the LB task. Lastly, these EF abilities are also relevant given that they are also known to be related to false belief understanding which was the original comparison for the LB task. While EF includes a variety of neurocognitive skills, inhibitory control, working memory, and cognitive flexibility are a coherent grouping for the preschool years (Carlson, Zelazo, & Faja, 2019).

[2] Allen and Bickhard (2018) included two versions of the task: the original task (LB-perceptual) and an attempted extension of the task (LB-structural). In the current study we are using the original task (LB-perceptual).

1.3.1. Inhibition

It is possible that the 3-step structure of the LB task creates inhibitory demands such that children in the third/reflection step were biased to perseverate (i.e., give the incorrect answer that the blocks would "fall"). Further, conflict inhibition is known to be related to false-belief tasks and so perhaps conflict inhibition is also responsible for the relationship between false-belief and LB tasks found by Allen and Bickhard. To further motivate the possibility that inhibition is playing an important role in children's LB performance, it should be noted that, in the original procedure, children who answered either of the first two steps incorrectly (i.e., they said that the single block would stay up) were prompted by repeating the question. Lastly, the falling affordance of the single blocks in the first two steps was made salient by doing a quarter rotation for each block, from vertical to horizontal and back to vertical, before starting the trial. This was done to reduce the relative novelty of the anticipations that would be needed to answer the first two questions (i.e., this action would increase the affordance of the object as "fall-able"). Thus, all of these features of the original procedure may have increased the inhibitory demands for correctly answering the question in the third/reflection step of the task.

1.3.2. Working memory

The *memory* demands in the LB task seem to be minimal. Children are looking at the two blocks and the questions by the experimenter are short and clear. However, *working memory* has also been modeled in terms of complexity (Case, 1992; Fischer, 1980; Pascual-Leone, 1970). Along these lines, Halford et al. (1998) have provided a well-developed working memory approach to development that is domain-general and quantitative/continuous. According to Halford and his colleagues' account, the age 4 transition and the development of meta-representational abilities involved with false-belief tasks can be understood in terms of relational complexity. Specifically, understanding one's own and others' beliefs can be construed in terms of binary relations in which children must grasp the relation between a percept and an object (e.g., Believed-location$_1$ (<object$_1$-location>, <percept>)). However, even if meta-representational abilities can be achieved through binary relations, the classic theory of mind tasks (e.g., false belief, appearance-reality, and perspective-taking) require ternary relations (Halford, Wilson, Andrews, & Phillips, 2014). For example, a relation with three variables can be built as "Find-object$_1$ (<known-event>, <actual-location>, <believed-location>)" to pass the false belief task (Halford et al., 1998). In the case of the change of location task, children not only need to think about the object's displacement from a box to a basket and the false belief about the object's location in the box, but also the object's actual location which requires understanding a relation between three variables (Halford et al., 2007).

However, the LB task would seem to involve binary relations (i.e., movement (<object$_1$>, <object$_2$>)) which are typically available to children by 2-years of age. Further, if the LB task requires only binary relations, then the correlations between LB and the FB tasks found by Allen and Bickhard (2018) are difficult to explain. Despite this a priori analysis, WM is an important component of EF and is thought to also involve important developments around age 4. Further, some of the core features of WM (e.g., mental rotation) are thought to be accomplished by reflection. Accordingly, it is worth testing empirically whether a typical WM task can explain performance on the LB task.

1.3.3. Cognitive flexibility

According to Zelazo (2004), children around age 4 are able to engage in a "new" type of reflection. This type of reflection enables children to construct hierarchically embedded rules that can be used to coordinate conflicting judgments in a given situation as measured by the Dimensional Change Card Sort task (DCCS, Frye, Zelazo, & Palfai, 1995; Zelazo, 2006). This difficulty in coordinating conflicting judgments has been used to explain performance on false-belief tasks as well (Benson, Sabbagh, Carlson, & Zelazo, 2013). Specifically, the child must coordinate their representation of an objects' location/identity with that of someone else. The conflict arises because it is the same object being represented from the two conflicting perspectives.

The point of hierarchical embedding is to create meta-rules that allow for real-time switching between judgments/perspectives and in so doing, to avoid any apparent conflict. For the DCCS, the meta-rule (if color game, sort by color; if shape game, sort by shape) reconciles the inability to play both games simultaneously. For the change of location FB task, the meta-rule (if my perspective, the box; if your perspective, the basket) also reconciles the inability of the marble to be in both locations simultaneously. For the LB task, there are not two situations with conflicting judgments so much as there are two competing affordances that need coordination. That is, the default perspective might be that both leaning blocks will fall, as predicted in response to the first two questions about the individual blocks, and a higher-order rule/perspective (i.e., reflection) might be needed to adopt the conflicting and less-salient perspective that the two blocks can be mutually supportive (i.e., not fall down).

1.4. The current study

The current study builds on previous work by Allen and Bickhard (2018) by: first, replicating the basic findings on the LB task; second, to test for alternative domain-general interpretations of performance on the LB task; and third, extending the assessment of reflection to two additional tasks. The analytic strategy follows from Allen and Bickhard which was itself based on work by Philip Zelazo and Joseph Perner who aimed to provide evidence for a domain-general shift around age 4. Specifically, these researchers designed tasks in different domains and looked at their relations before and after controlling for age.

Frye et al. (1995) have argued that age-partialled correlations remove the basic effect of development on performance. Therefore, if tasks in two different domains both rely on some domain-general development, then they should be correlated before but not after controlling for age. However, if the two tasks have additional and overlapping performance demands (or they are relatively pure measures of the underlying ability), then they may be correlated before and after controlling for age. Accordingly, it was predicted that the new reflection measures would show: 1.) an age 4 transition; 2.) correlations before and after controlling for age; 3.) remain correlated after also controlling for EF. In contrast, if aspects of EF are important for explaining performance on LB, then EF would be expected to be correlated with the reflection measures after controlling for age.

2. Method

2.1. Participants

Based on Allen and Bickard's (2018) study, we collected data from seventy-two (38 male, 34 female) Turkish middle and upper-middle class (3-year-olds, $M = 41$, $SD = 3.7$; 4-year-olds $M = 54$, $SD = 3.3$; 5-year-olds, $M = 64$, $SD = 2.8$). Power analysis with a medium to large effect size ($r = 0.40$) and a power of 0.80 to 0.90 resulted in a sample size of 46−75. Five additional children were tested but excluded from the study altogether (2 were below 36 months, 2 were over 72 months, and one had missing age data). The majority of children participated in the study at their preschools in a metropolitan city, Istanbul. Istanbul is the largest city in Turkey with a population of approximately 15 million residents. Before the study began, Institutional Review Board approval was obtained from Bilkent University and permission was granted from the Ministry of Education of Turkey.

2.2. Target measures

2.2.1. Leaning blocks (LB) task

This task measures children's ability to reason explicitly about the relations between objects (Allen & Bickhard, 2018). Children were asked to predict the outcome of novel manipulations done with the blocks. Each trial of the task included two single object questions and one object relation question. Three trials were used with different types of blocks (small wooden blocks, large wooden blocks and two large rectangular pieces of cardboard). In each trial, the first block was held in a pincer grip from one of its edges at an inclined position. Children were asked, "What will the block do if I let go? Will it fall or stay-up?". After the block was released, the block was removed from the child's sight and the second single-object question was repeated with the second block held on the opposite side of the first. After the single-object questions, the two blocks were held in pincer grips such that they were leaning against each other at a 45 degree angle and the experimenter asked, "What will the blocks do if I let go? Will they fall or stay-up?". All questions were only asked once (i.e., no prompts) and no warm-up with the blocks was given. In the second and third trials, the procedure was repeated with two large blocks and then with two large rectangular pieces of cardboard. Children received a score of one for each answer for a total score out of 9 (3 possible points for each set of blocks).[3]

2.2.2. Candy monster (CM) task

A second (*kinetic*) reflection task was designed to assess children's ability to explicitly reason about objects. In this task, children had to mentally transform the order of three colored balls being inserted into a tube (i.e., they had to indicate the opposite order in which the balls were inserted). To minimize language and working memory demands, the tube was transparent and children were familiarized with a scenario about a monster who likes to eat candies. To start the CM task, children were introduced to a brown plush monster who loves eating candy. Next, children were shown a transparent tube with openings at each end but one of the openings had a serrated edge (see Fig. 1). The tube was placed horizontally and children were told, "Look, this is the tube from which the monster loves to eat candies!". Then three sand-filled balloons of different colors were introduced to the child as monster-candies and two of them were inserted into the tube from the non-serrated opening. Children were told, "If you are ready, I will turn the tube to let the monster eat the candies!". When the child said that s/he was ready, the experimenter rotated the non-serrated edge of the tube toward the monster, moved the monster toward the candies, and made some eating gestures. This familiarization trial was repeated once again by replacing one of the candies inside the tube with the third candy that had been introduced in the beginning.

For the test trial, the same three candies were inserted into the tube one by one. The order was such that the last candy inserted, which the monster would eat first, was always a different color from the first candy the monster ate in the familiarization trials (i.e., if they used their same answer from familiarization, they would be wrong). Different from the familiarization trials, the tube was not turned around. Instead, it was explained to children that the serrated opening would hurt the candy monster, "Look, this end of the tube is broken and has sharp edges on it. If the monster eats its candies from this 'sharp' end, it will hurt and his mouth will also be cut!". The experimenter indicated that they would turn the tube (i.e., an appropriate gesture was done by the experimenter's hand to show the child that other end of the tube would be turned to let the monster eat its candies from the unbroken end). Then children were asked "Which candy will the monster eat first, second and third after I turn the tube?". Since children either answered all three questions correctly or only one question (i.e., there was no child that correctly answered two of three questions), the task was coded as pass (correctly answering all three questions) or fail (correctly answering one question) for a range from 0−1.

[3] Allen and Bickhard (2018) used different criteria for coding such that children needed to correctly answer the first two questions for their response on the third question to count. However, that coding assumes that the first two questions do not also involve whatever processes are relevant for the third question (i.e., reflection). Since the purpose of the current study is to test alternative interpretations, we used a more inclusive, and theory neutral, coding that includes all of the child's responses. For the basic set of relations between variables, the original scoring was analysed and any differences noted (see Table 2).

Fig. 1. Candy Monster (CM) Task: balloons looking like different colors of candies, the brown plush monster and the transparent tube with a serrated edge (left).

2.2.3. Elastic band (EB) task

The EB task was used to further elucidate the nature of preschoolers' reflective abilities. The original EB task was developed by Piaget and was used to assess *transformational* imagery (Piaget & Inhelder, 1971, pp. 215-218). The task has been modified in order to be used with preschoolers. For the familiarization phase, children were shown a wooden plate with four holes in the shape of a square (see Fig. 2). Next, four blue pegs were inserted into the holes and children were told "Look, these pegs can be put in the holes and we can take them out". Next, children were shown a black elastic band and told, "This elastic band is so flexible and it can look like many shapes" (The band was stretched by the experimenter's hands in different ways). Next, children were shown four cartoonish pictures of the elastic band around the pegs in different configurations and told, "These black figures are pictures of the elastic band in different shapes and these dots in the pictures are pegs" (see Fig. 2).

The EB task had one familiarization question and two versions. In the familiarization question, after the pegs were placed in the holes, children were asked, "If I place this elastic band around the pegs which picture will it look like?". If children could not give any answer or gave an incorrect answer, the experimenter again familiarized children with the materials up to three times. Only four children (two 3-year-olds and two 4-year-olds) could not give a correct answer within the three trials. After the third trial, the elastic band was put around the pegs to show them that the elastic band took the shape of a square.

In the first version (EB 1), children were then shown four different pictures of the elastic band wrapped around three pegs. These were all triangles in different orientations (see Fig. 2). Then children were redirected back to the elastic wrapped around the four pegs and told, "If I take out this peg, which shape, from these pictures, will the elastic band take?". The question was repeated four times for each peg in the plate. In the second version, called EB 2, children were shown four pictures of the elastic band in the shape of a straight line (i.e., two pegs removed) but in four different orientations (see Fig. 2). Again, children were told, "If I take out these two pegs, which shape, from these pictures, will the band take?". The question was repeated four times: two diagonal, one vertical, and one horizontal shape with appropriate gestures of a pegs' removal. The two versions, EB 1 and EB 2, were scored separately out of 4 (one point for each question).

2.3. Control measures

2.3.1. Day-Night task

The Day-Night Stroop task was used to assess preschoolers' inhibitory control (Gerstadt, Hong, & Diamond, 1994). In this task, children are required to label a picture of the moon as "day" and a picture of the sun as "night". Before the test trials, children needed to correctly identify two cards in a row (correctly according to the rule). After the familiarization trials, sun and moon pictures were shown to children one at a time in a pseudo random order for a total of 16 trials (score ranged from 0–16).

2.3.2. Balloon task

A new task was developed to assess children's working memory ability and is called the Balloon task. In the Balloon task, children were first shown an opaque tube, held in a vertical position. The experimenter also let children look inside the tube to see that it was empty and that balloons could only fit on top of each other. Next, children were shown three balloons whose colors were blue, green, and red. Then children were told to watch carefully as the experimenter put the balloons in the tube one by one. After confirming that they understood what the experimenter told them, the blue balloon was put into the tube with the following commentary "First, I put the blue balloon", then, the red balloon was put into the tube and the experimenter said "After that, I put the red balloon"; finally, the green balloon was put in the tube and the experimenter said "Lastly, I put the green balloon". The height of the tube was sufficient such that children could not see the green balloon.

For the test phase, children were asked, "If I want to pull out the balloons one by one from the top of the tube, which color will come out first?". After the child gave the answer for the first balloon, the experimenter asked the same question for the second and third balloons separately. The measure included two trials where the same task was repeated with different colored balloons (i.e., white, black, and yellow). Children's performance was assessed out of 6 (they received one point for each correct answer in a given trial).

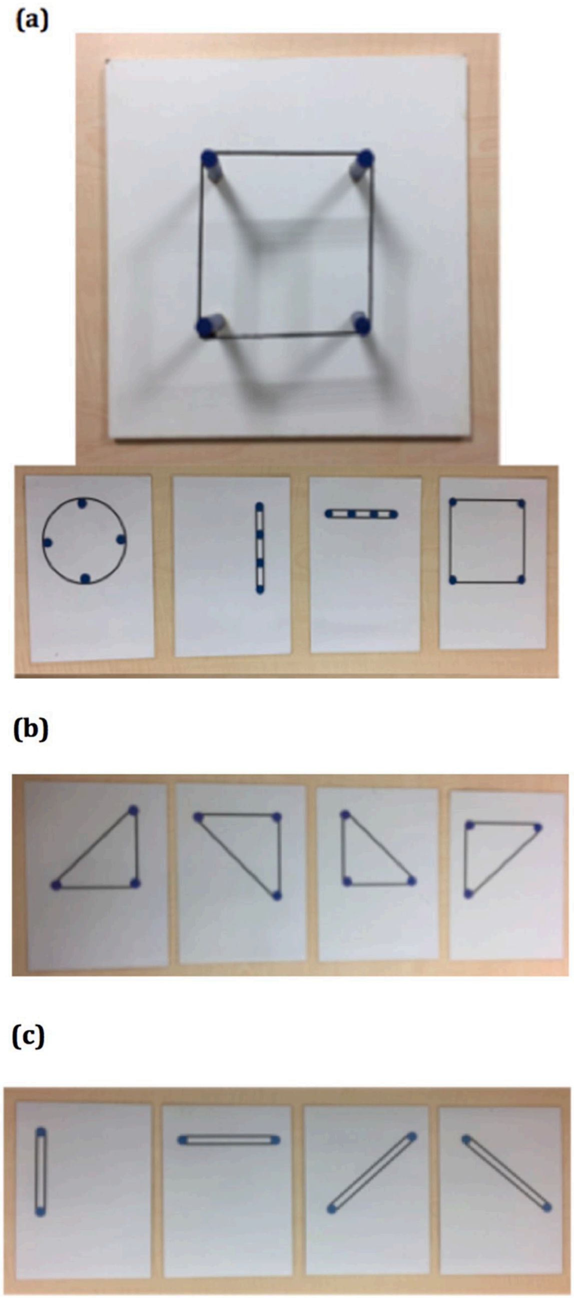

Fig. 2. Elastic Band (EB) Task: (a) the wooden plate with 4 pegs, surrounded by the elastic band and the possible options for the familiarization question; (b) the possible options for the first version (EB1); (c) the possible options for the second version (EB2).

2.3.3. Dimensional change card sort (DCCS) task

The DCCS was used to assess preschoolers' executive functioning (Zelazo, 2006). In this version of the task, there are three consecutive phases: "shape game", "color game", and "border game", involving blue and red pictures of a car or an elephant. For the color game, children sorted picture cards according to color. For the shape game, children needed to switch and sort the same set of cards by shape. For the border game, children needed to switch between the two games in accordance with a cue on the border of the cards. Children only progressed to subsequent games if they passed the prior game (i.e., scored 5 out of 6 correct responses for the first two games). The border game was out of 12 and required a score of 9 for children to be considered as having passed. For the total EF score children received 1 point for passing each game (max score 3).

2.4. Procedure

Each child completed three reflection tasks (Leaning Blocks (LB), Candy Monster (CM), and Elastic Band (EB)) and three control tasks (the Balloon task for working memory, the Dimensional Change Card Sort task (DCCS) for executive functioning and the Day-Night task for inhibition). Children participated in a small and quiet room. The whole experiment took about 30 min. The experiment was split into two phases. In phase one, children were given the Day-Night, EB, and LB tasks in that order. After the LB task, the child and experimenter took a break by building a castle from the blocks used in the LB task. The break took approximately 3−5 min. After the break children were given the DCCS, Balloon, and CM tasks, respectively.

3. Results

The three aims of the current study were: 1.) to *replicate* the age 4 transition found for the LB task by Allen and Bickhard (2018); 2.) to *test* for alternative domain-general interpretations of those results; 3.) to *extend* the findings for LB to other "reflection" tasks.

3.1. Age analyses for LB, CM, and EB

The data for two children were absent for the LB task (n = 70) and could not be included in the analyses (one child did not want to play and the other child's answers were unclear such that the coders could not reach an agreement on the response). Four children had missing data on the CM task (n = 68) such that three children did not want to play the whole game and one child played only the familiarization part. One child had missing data for the EB task (n = 71) because they did not want to play.

Table 1 shows the performance of children for the LB, CM, and EB tasks. The two sub-tasks of the EB task were not correlated after controlling for Age. This suggests at least two possibilities. First, it may be the case that the two sub-tasks measure two different abilities. Second, the two sub-tasks may measure the same ability but one of them may have additional performance factors that are strongly correlated with age. Regardless of the interpretation, the lack of correlation led to treating the two versions separately (i.e., EB1 & EB2).

Table 1
Means and Standard Deviations for all Tasks by Age.

		Target Variables				Control Variables		
Age		Leaning Blocks (0-9)	Candy Monster (0-1)	Elastic Band 1 (0-4)	Elastic Band 2 (0-4)	Balloon (0-6)	DCCS (0-3)	Day-Night (0-16)
3-year-olds	Mean	6.2 (2.1)	0.4 (0.5)	1.1 (0.9)	0.7 (0.7)	2.2 (1.6)	1.0 (0.6)	11.4 (4.4)
	N	25	23	25	25	24	26	24
4-year-olds	Mean	8.1 (1.1)	0.9 (0.4)	1.5 (1.4)	2.2 (1.5)	3.8 (2.1)	1.7 (0.7)	13.1 (3.5)
	N	26	26	26	26	26	26	26
5-year-olds	Mean	8.5 (0.7)	0.9 (0.3)	2.5 (1.4)	3.0 (1.3)	4.7 (1.5)	2.0 (0.4)	14.5 (2.9)
	N	19	19	20	20	19	19	20
Total	Mean	7.5 (1.8)	0.7 (0.5)	1.6 (1.4)	1.9 (1.5)	3.5 (2.0)	1.5 (0.7)	12.9 (3.9)
	N	70	68	71	71	69	71	70

Note: Standard Deviations for Scores are given in parenthesis.

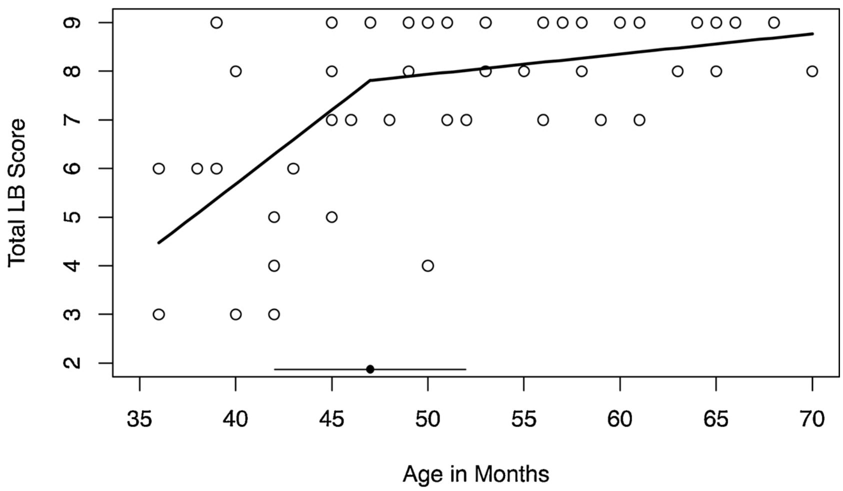

Fig. 3. Scatterplot of the breakpoint model for the Leaning Blocks task (LB; out of 9). The horizontal line at the bottom indicates the 95 % CI for the breakpoint.

A one-way ANOVA was conducted for LB with Age as the independent variable. Analysis indicated a main effect of Age, $F(2, 67) = 16.14$, $p < .001$, $\eta_p^2 = .33$. LSD follow up analysis showed that 3-year-olds differed from 4-year-olds, $p < .001$, and 5-year-olds, $p < .001$, but that 4-year-olds did not differ from 5-year-olds, $p = .423$. In order to follow-up on the idea of a threshold effect related to the age 4 transition on the LB task, a piecewise linear regression analysis was conducted in R using the segmented package (Muggeo, 2008). An initial guess parameter was used for the breakpoint of $\psi = 45$ months of age and $k = 10$ segments. According to the Davies test, the breakpoint was statistically significant at 46.67 months of age for a change in the slope, $p = .006$, 95 % CI = 42.03–51.96 months of age. The bivariate correlations between children's age and their performances on the LB task were $r = .51$, $p = .011$, below the breakpoint and $r = .24$, $p = .115$, above it (see Fig. 3 for the breakpoint model).

For the CM task, a Chi-square analysis by Age showed a pattern of results similar to what was found for LB, χ^2 (n = 68) = 14.23, $p < .001$. Specifically, 3-year-olds differed from 4-year-olds, χ^2 (n = 49) = 9.12, $p = .003$, and 5-year-olds, χ^2 (n = 42) = 9.59, $p = .002$; however, 4-year-olds did not differ from 5-year-olds, χ^2 (n = 45) = 0.22, $p = .636$. For the EB1 and EB2 tasks, separate one-way ANOVAs were conducted. For EB1, results indicated a main effect of Age, $F(2, 68) = 6.76$, $p = .002$, $\eta_p^2 = .17$. LSD post-hoc tests showed the opposite pattern for what was found with the LB and CM tasks. That is, for the EB1 task, 5-year-olds differed from 4-, $p = .011$, and 3-year-olds, $p < .001$; but, 4-year-olds did not differ from 3-year-olds, $p = .286$. For the EB2 task, there was also a main effect of Age $F(2, 68) = 19.92$, $p < .001$, $\eta_p^2 = .37$. LSD post-hoc tests revealed pair-wise Age differences between all three groups. That is, 3-year-olds differed from 4-, $p < .001$, and 5-year-olds, $p < .001$; and, 4-year-olds also differed from 5-year-olds, $p < .034$.

3.2. Simple and age-partialled correlations amongst the tasks

Simple correlations for all tasks can be found in Table 2. All tasks were strongly correlated with Age, $r = .41$–$.63$ (see Fig. 3 for the scatter plot between Age and the LB task). Most of the target tasks (LB, CM, EB1, & EB 2) were correlated with most of the control measures (Day-Night, DCCS, & Balloon tasks, see Table 2). LB was correlated with the other three target measures (for CM, $r = .46$; for EB1 and EB2, $r = .29$ and .27 respectively). LB was correlated with two of the three control measures (for Day-Night, $r = .39$; for DCCS, $r = .49$).

Age-partialled correlations for all tasks can be found in Table 3. Only four correlations were significant at the .05 level after controlling for Age. Of particular relevance is the continued correlation between LB and CM, $r = .25$, and the lack of correlations between LB and Balloon, DCCS, and Day-Night respectively. Lastly, EB2 but not EB1 was significantly or marginally correlated with all three control measures after controlling for Age. Finally, after partialing for Age and the three control variables, there was only one correlation between target measures. Specifically, LB continued to be correlated with the CM task (see Table 4).

Table 2
[a]Simple Correlations Among the Tasks.

Variable	Age in Months	LB	CM	EB 1	EB 2	Balloon	DCCS
Leaning Blocks	0.62***						
Candy Monster	0.47***	0.46***					
Elastic Band 1	0.41***	0.29*	0.09				
Elastic Band 2	0.58***	0.27*	0.34**	0.32**			
Balloon	0.50***	0.19	0.47***	0.26*	0.53***		
DCCS	0.63***	0.49***	0.37**	0.36**	0.52***	0.43***	
Day-Night	0.41**	0.39***	0.33**	0.12	0.08	0.29*	0.37**

[a] Note that when the analysis was done with Allen and Bickhard's (2018) criteria (i.e., when children's performance was coded out of 3), two correlations became non-significant (the correlation between LB and EB 1 ($r = .20$, $p = .103$) and LB and EB 2 ($r = .20$, $p = .100$).

Table 3
Age Partialled Correlations Among the Tasks.

Variable	LB	CM	EB 1	EB 2	Balloon	DCCS
Leaning Blocks						
Candy Monster	0.25*					
Elastic Band 1	0.02	−0.16				
Elastic Band 2	−0.12	0.10	0.12			
Balloon	−0.16	0.31*	0.04	0.35**		
DCCS	0.17	0.13	0.13	0.26*	0.17	
Day-Night	0.16	0.15	−0.06	−0.21[t]	0.07	0.11

*$p < .05$, *$p < .01$ ***$p < .001$, [t]$p < .10$.
Note: Age was partialled in months.

Table 4
Age, Balloon, DCCS, Day-Night Partialled Correlations.

Variable	LB	CM	EB 1
Leaning Blocks			
Candy Monster	.29*		
Elastic Band 1	.04	−.18	
Elastic Band 2	−.10	.01	.09

*$p < .05$, **$p < .01$ ***$p < .001$, $^{t}p < .10$.
Note: Age was partialled in months.

4. Discussion

There is growing evidence to support the idea that there is a domain-general development around age 4 that is responsible for the changes in preschoolers' performance on a variety of tasks within and across domains (Halford et al., 1998; Perner, 1991; Zelazo, 2004). The interactivist model of knowing levels proposes reflection as being responsible for that domain-general development (Bickhard, 1980, 1992; Campbell & Bickhard, 1986). The leaning blocks task was designed to differentiate between reflective and pre-reflective object representation and the results from Allen and Bickhard (2018) provide evidence for the interactivist model. The current study sought: to *replicate* the original findings for the LB task; to *test* for alternative explanations for performance on LB; and, to *extend* the findings from LB to two new tasks.

4.1. Does LB performance replicate in a different cultural sample?

Age analyses from the current study replicated the threshold effect for LB that was found by Allen and Bickhard (2018). That is, performance jumped between ages 3 and 4 but remained relatively constant after age 4. The reason for the threshold effect is thought to be due to the minimal performance factors needed to pass the task. Reflection is argued to enable the possibility of transcending the thought-in-action constraint of interactive knowing (i.e., the possibility of representing the object qua object). Once this epistemic capacity is in place, passing the task becomes "trivial". Accordingly, the LB task is argued to be a relatively pure measure of reflection. Some indirect evidence for this conclusion may also come from the fact that children did receive feedback after each trial (i.e., the blocks were released). Although some children switched to correct responding after their original failure, over two thirds continued to fail the task. This suggests that some sort of insight is needed for "learning" to take place, and, the current argument is that such insight is a consequence of the development of reflection.

Three differences from the original study by Allen and Bickhard (2018) further suggest that the replication of the age effects is robust. First, the current sample was Turkish in contrast to the American sample used in the original study. With respect to family structure and parenting, Turkey is more similar to the emotional interdependence aspect common in collectivist cultures from the "East" than to that of the "West" (Göregenli, 1997; Kagitcibasi, 1996; Kagitcibasi & Ataca, 2005). The lack of cross-cultural variance is consistent with the underlying assumption of interactivism that the age 4 transition is due to brain maturational developments. Second, age in the current sample was spread across the range from three to five. In the study by Allen and Bickhard, children were clustered within plus or minus two months of ages three, four, and five. It might be argued that such a clustering biased a categorical interpretation of the results from the original study. Third, the procedure and scoring in the current study was more theory-neutral. That is, the procedure did not cue/prompt children on the first two questions about what would happen to the single blocks when released (i.e., treat them as control questions). Accordingly, the current scoring used children's responses to all of the questions for a total out of nine. This also allowed for a more direct test of the assumed threshold age effect through the piecewise regression analysis. While the change in procedure did affect performance on the first two questions[4] it did not change the overall pattern of performance across ages.

4.2. What is the relationship between reflection and other domain-general processes?

An a priori task analysis of leaning blocks suggests relatively few performance factors: the instruction and response are straightforward and concise and the visual presence of the blocks minimizes memory demands. That said, the second aim of the current study was to test for potential EF performance factors empirically. While all three EF abilities were correlated with LB, the correlations were not present after controlling for age. This suggests that the EF measures assess something different from the reflection abilities needed for the LB task. Further, the second "pure" reflection task, candy-monster, was correlated with LB after controlling for age and the three EF factors. This suggests that these tasks may both be relatively pure measures of reflection; however, it is also possible that they share additional task demand not relevant to the EF tasks. The quite different procedures and materials of the LB and CM tasks would favor the first interpretation. In contrast to the CM task, correlations were not present for the LB and EB tasks after controlling for age nor after controlling for age and EF. Collectively, these results suggest that the EB tasks do not directly measures reflection and are likely to

[4] In the current study, 36 % of children answered one or more of the first two questions incorrectly; while in Allen and Bickhard (2018), only 16 % of children answered one or more of the first two questions incorrectly.

involve working memory demands (more on this below).

That performance on the LB tasks involves an age 4 transition and that the EF control tasks (Sun-Moon, Balloon, DCCS) were not related after controlling for age provides direct evidence for the interactivist proposal about the knowing levels of model of implicit versus explicit knowing. The relevance of these findings for broader ideas about the development of EF abilities is less clear. From the current perspective of interactivism, inhibitory control, working memory, and cognitive flexibility at age 4 all *require* reflection; though reflection alone is not *sufficient*. However, this leaves open how to explain the seemingly similar forms of inhibitory control and working memory prior to age 4 (cognitive flexibility is already assumed to undergo a qualitative change at age 4; Zelazo, 2004). From the current perspective, such earlier forms would need to involve different processes than reflection. Such a position would make sense if the presumed unity of inhibitory control and working is more a matter of appearance than underlying process.

As has been noted by others, inhibition is difficult to understand as a unitary phenomenon because there is no clear reason how the inhibitory demands on different tasks and at different ages are united (Müller & Kerns, 2015; Zelazo & Müller, 2011). While there may be some general properties related to what gets called inhibitory control, there would also seem to be some content specific aspects such that some inhibitory problems are also a consequence of insufficient ability or knowledge. That is, inhibition is, at least sometimes, a property of being a novice. Consider learning to play basketball. Let us assume that all adult players have similar general inhibitory ability but the novice players will, undoubtedly, fail to inhibit the tendency to pass the ball even after they realize it is going to be intercepted. Presumably, what changes with expertise, is not inhibition as a content-independent ability, but the skill/knowledge about how to play basketball and as a consequence the expert player engages in adequate "inhibition" so as not to give the ball away.

For working memory too, the implication of reflection as enabling all EF abilities around age 4 is that earlier examples of working memory are qualitatively different. This is in contrast to most complexity models of working memory that explain development in terms of an increasing resource involving quantitative development (Case, 1992; Fischer, 1980; Pascual-Leone, 1970). For Halford et al. (1998), the increase was in terms of the number of possible relations. However, the age effects on LB are not clearly explainable by the relational complexity approach and the lack of an empirical connection between LB and the modified backward digit-span task (i.e, Balloon task) reinforces that conclusion. That said, the concept of working memory is broader than just relations and backward digit-span is not the only way to measure working memory. Further, interactivism would need to explain how implicit presupposition is able to account for unary and binary relations in order to pose a direct challenge to complexity theory.

The DCCS task is similar to LB in that both tasks are tightly connected to an underlying theory that motivated their construction. Further, the levels of consciousness model (Zelazo, 2004) has some conceptual overlap with levels of knowing in terms of function. That is, both models suggest that reflection accounts for qualitative developments in children's abilities. The explanation of LB performance in terms of rule complexity was more clear than for relational complexity and the DCCS task is a more direct measure of rule complexity than is backward digit-span a measure of relational complexity. Accordingly, the lack of statistical relation between the DCCS and LB task implies a more clear difference between levels of consciousness and levels of knowing. One of the main differences between the two frameworks has to do with the underlying models of representation and cognition. For levels of consciousness, reflection is modeled in terms of recurrent processes that re-processes information. For levels of knowing, reflection is an abstraction from implicit content in the organization of interactive knowing. Future research may explore more closely the theoretical and empirical differences between the two models.

In summary, the action-based perspective within which interactivism is situated assumes that all of these domain-general EF abilities are content-specific in some constitutive sense. This assumption is in contrast to the information processing perspective within which most EF models are situated. For interactivism, reflection is an *enabling* constraint that makes a qualitatively new way of knowing the world possible but particular internal and external interaction is still needed to develop the various abilities captured by the different EF tasks (See Doebel, 2020, for a developed proposal that argues for the need to reconceptualize EF along more content-dependent lines).

4.3. What are the developmental patterns for performance on the new reflection tasks?

Candy-monster and the EB tasks were designed to further explore the LB task as related to reflection. The CM task was intended as a second "pure" measure of reflection while the EB tasks were intended to explore possible relations between the two types of anticipatory imagery (*kinetic* and *transformational*, (Piaget & Inhelder, 1971, p. 352)). The LB task is kinetic because the reasoning does not involve a transformation of the blocks themselves. Similarly, the CM task was designed to be solvable by the reflective child without the need to engage in a transformation of the tube (i.e., by reasoning that they could look to the end opposite the monster and then answer in sequence). Again, novelty is essential for this task to differentiate reflective and pre-reflective reasoning given that they could also just learn this strategy through sufficient experience with adequate feedback.

Two of the results suggest that the CM task is a kinetic-anticipatory task similar to LB. First, the age effects for CM were similar to those for LB such that children showed the threshold jump at around age 4. Second, LB and CM were correlated after controlling for age and continue to be correlated after also controlling for the EF variables. However, unlike LB, the CM task was correlated with the Balloon tasks after controlling for age. There are at least two possible ways that children could solve the CM task: first, children may use a direct inference (i.e., look to the opposite end of the tube and answer in sequence); or, second, use a mental rotation of sorts (mentally rotate the tube/puppet and answer in sequence). Given that working memory has been characterized by abilities like mental rotation, the relation between CM and the Balloon task may suggest that latter solution. However, it is also possible that the correlation between CM and Balloon tasks was more of an artifact due to their similarity as tasks. The two tasks share in terms of their appearance (i.e., balloons and tubes) and what children need to do (i.e., report the correct sequence). Future research could include additional WM tasks to better understand specifically how reflection is being used by children to solve the CM task. For example, some other "preschool friendly" working memory tasks have been developed that are not similar to the CM task (e.g., a spatial span task, Müller, Liebermann-Finestone, Carpendale, Hammond, & Bibok, 2012; counting and labeling task, Carlson, Moses, & Breton, 2002).

Performance on the EB tasks provided some unexpected results suggesting that they should be interpreted with caution. The overall patterns of results for the two EB tasks suggests that they do not primarily measure reflection and that they may not measure the same underlying ability. First, it was expected that EB1 and 2 would be correlated with each other before and after controlling for age but this was only the case for the former correlation. The lack of correlation after controling for age suggests at least two possibilities. It may be that the two versions of the task do not measure the same ability or that one of the tasks may have additional performance factors that are themselves correlated with age. The second reason the EB tasks are unlikely to measure reflection is that the *assumed* difficulty of the two tasks was not found empirically. While it was assumed that the removal of two pegs for the EB2 task would be harder than the removal of one peg, this was not the case. The reasoning for the a priori assumption was that 2 pegs would require *more* change to the shape of the elastic; however, the end-point for EB2 was simpler than for EB1 (i.e., a straight line versus a triangle). It is possible that children did not engage in a mental transformation (i.e., that they did not reason about the change of the elastic band from a square to a triangle/straight line). Instead, they might only have used a lower-level "perceptual" strategy focused on the end-state that the elastic band would take. Specifically, when the experimenter asked children about what would happen after removing the two pegs, they were simultaneously touching both of them. This raises the possibility that children could more easily scan the remaining two pegs and in so doing would be looking at the answer (note also that the correct end-state was one of the 4 options presented to them with the flash cards). This lower-level strategy may have some support in terms of the relations between EB2 and the Balloon and DCCS tasks. While shifting between the two remaining pegs, children would need to mentally represent the line and coordinate that representation with the correct answer from the pictures.

Future research will be needed to delineate the possible processes involved for the EB tasks. What is more clear is that the EB tasks as currently constructed do not seem to be adequate measures of the development of reflection. There were no age 4 threshold effects for either task and neither task was correlated with LB or CM after controlling for age. Finally, EB1 and EB2 showed different relations to the EF control tasks such that EB2 may be more similar to other EF tasks and EB1 may be too difficult for the younger age group.

4.4. Implications for physical and social cognition

The development of reflection and the modeling of implicit versus explicit knowing may contribute to some of the well-known debates related to children's physical and social cognition before age 4. In terms of physical cognition, interactivism provides an alternative anticipatory explanation for the "striking" decalages found by nativists across infancy for "representing" objects in *occlusion* events but not *containment* or *covering* events (Allen & Bickhard, 2013; Baillargeon, 2008). After about two years of age, toddler behavior is generally not going to differentiate between implicit versus explicit object knowledge. The LB task is one of the few physical reasoning tasks in which the implicit versus explicit representing of object makes a difference for behavior. More broadly, the interactivist account of implicit content is one of the few models with adequate specificity to fulfill the promise of action-based approaches to account for the emergence and development of new forms of representation.

In terms of social-cognition there is greater support for the possibility that, prior to age 4, children cannot represent any mental states per se. For more empiricist versions of this position, behavioral rules are thought to underlie the social complexity found for young children and other social species (Heyes, 2014; Povinelli & Vonk, 2003). However, behavioral rules are still explicit, they just do not include mental state content. More constructivist accounts have also been proposed to explain infant and toddler performances through some sort of anticipatory process (Carpendale & Lewis, 2015; Fenici, 2015); however, there is still no account of implicit content.

4.5. Limitations and future directions

In order to further explore their reasoning on the LB task, children might be asked to justify their predictions. As it has been observed in the EF literature (Morton, Trehub, & Zelazo, 2003), there might be a positive relation between children's performance on the LB task and their explanations. Also, a wider variety of EF measure could be used to further test whether performance on the LB task can be explained by EF. In particular, the reliance of the DCCS on rules may make it unlike other measures of cognitive flexibility. Further, given the brain maturation thought to underlie the development of reflection, future research may try to use a micro-genetic design with measures that involve more variance and a restricted age range (3.5-4-years perhaps). Such a design may also allow for more sophisticated statistical analyses to tease apart alternative explanations. Future work could also clarify the working memory demands for the CM task by including alternative measures. In particular, a working-memory task that focuses on mental rotation may be helpful. Second, the candy monster and tube were placed horizontally in front of the child. This may have required some minimal perspective taking ability that could be eliminated by placing the task materials in-line with the child or could be tested empirically by looking at the relation between a preschool-aged perspective-taking task (e.g., Flavell, Everett, Croft, & Flavell, 1981) and the CM task. Additionally, it was discussed that children may have solved the EB2 task without taking the modification of the elastic into account (i.e., disregarding the changes in the elastic band's shape by only reasoning about the end-state). Further, EB1 may have been too difficult for this age group. Accordingly, changes should be made to simplify the task and ensure that children are required to reason about the transformation. The original procedure from Piaget did not involve flashcards but also seemed too difficult for the preschool age range. Thus, more work is needed to make the task appropriate for both the age group and the presumed ability being measured.

4.6. Conclusion

From the interactivist perspective, pre-reflective children are restricted to the thought-in-action constraint before age 4 (i.e., they can only know the world through their interactions with it). The current study provides direct evidence for the idea that without reflection, objects cannot be explicitly represented, and therefore, without reflection children fail tasks that require reasoning about object relations. Further, this limitation regarding relations applies even more clearly to the perceptually unavailable properties of people (e.g., mental states). This means that children cannot engage in the explicit representation of perceptually hidden mental states prior to the development of reflection. Accordingly, a major implication of this study is that pre-reflective children must be using interactive knowing to navigate both their physical and their social worlds. As a consequence, this implication undermines any theoretical approach to social cognition in which children are assumed to be engaged in mental-state attribution prior to age 4.

Cognitive science is replete with examples of a basic division between two categories of knowing (e.g., procedural-declarative; tacit-explicit; heuristic-rule-based; automatic-deliberative; system1-system2). Further, there are a number of proposals suggesting a domain-general transition around age 4. However, one of the unique benefits of the interactivist model is its rich ontology for modeling pre-reflective knowing and for the multiple sense of explicit and implicit representation (Bickhard, 1998). Further, this rich ontology provides internally consistent models with other areas such as language, perception, motivation, emotions, sociality, persons and so on (Bickhard, in-preparation; Bickhard & Terveen, 1995). Lastly, the interactivist model is convergent with our understanding of the natural world and has a coherent account of how the mind is manifest in the brain (Bickhard, 2015a, 2015b). Accordingly, the age 4 transition as explained by the interactivist model promises to be a fruitful avenue to pursue both theoretical and empirical issues related to qualitative change.

Sources of funding

None declared.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgments

We would like to thank all of children and their parents who participated in this study as well as the preschools for helping with data collection. We would also like to thank the anonymous reviewers for helpful comments and feedback. Lastly, a special thanks to Berfu Ulusoy for all the help with preparation of the materials and data collection.

References

Allen, J. W., & Bickhard, M. H. (2013). Stepping off the pendulum: Why only an action-based approach can transcend the nativist–empiricist debate. *Cognitive Development, 28*(2), 96–133. https://doi.org/10.1016/j.cogdev.2013.01.002

Allen, J. W., & Bickhard, M. H. (2018). Stage fright: Internal reflection as a domain general enabling constraint on the emergence of explicit thought. *Cognitive Development, 45*, 77–91. https://doi.org/10.1016/j.cogdev.2017.12.005

Atance, C. M. (2008). Future thinking in young children. *Current Directions in Psychological Science, 17*(4), 295–298.

Baillargeon, R. (2008). Innate ideas revisited for a principle of persistence in infants' physical reasoning. *Perspectives on Psychological Science, 3*(1), 2–13. https://doi.org/10.1111/j.1745-6916.2008.00056.x

Benson, J. E., Sabbagh, M. A., Carlson, S. M., & Zelazo, P. D. (2013). Individual differences in executive functioning predict preschoolers' improvement from theory-of-mind training. *Developmental Psychology, 49*(9), 1615. https://doi.org/10.1037/a0031056

Bickhard, M. H. (1980). A model of developmental and psychological processes. *Genetic Psychology Monographs, 102*, 61–116.

Bickhard, M. H. (1992). Commentary on the age 4 transition? *Human Development, 35*(3), 182–192.

Bickhard, M. H. (1998). Levels of representationality. *Journal of Experimental & Theoretical Artificial Intelligence, 10*, 179–215. https://doi.org/10.1080/095281398146824

Bickhard, M. H. (2021). *The whole person: Toward a naturalism of persons — Contributions to an ontological psychology*. in preparation.

Bickhard, M. H. (2009a). Interactivism: A manifesto. *New Ideas in Psychology, 27*(1), 85–95. https://doi.org/10.1016/j.newideapsych.2008.05.001

Bickhard, M. H. (2015a). Toward a model of functional brain processes II: Central nervous system functional macro-architecture. *Axiomathes, 25*(4), 377–407.

Bickhard, M. H. (2009b). The interactivist model. *Synthese, 166*(3), 547–591. https://doi.org/10.1007/s11229-008-9375-x

Bickhard, M. H. (2015b). Toward a model of functional brain processes I: Central nervous system functional micro-architecture. *Axiomathes, 25*(3), 217–238.

Bickhard, M. H., & Campbell, R. L. (1989). Interactivism and genetic epistemology. *Archives de Psychologie, 57*(221), 99–121.

Bickhard, M. H., & Terveen, L. (1995). *Foundational issues in artificial intelligence and cognitive science: Impasse and solution* (Vol. 109). Elsevier.

Campbell, R. L., & Bickhard, M. H. (1986). Knowing levels and developmental stages. *Contributions to human development*. Basel, Switzerland: Karger.

Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development: An International Journal of Research and Practice, 11*(2), 73–92. https://doi.org/10.1002/icd.298

Carlson, S. M., Zelazo, P. D., & Faja, S. (2019). De la-A ux-fonction. s executive. s: Une synthese neurocognitive et socioaffective en neurosciences cognitives et en psychologie du développement. *ANAE-Approche Neuropsychologique des Apprentissages chez l'Enfant, 31*(160), 371–413.

Carpendale, J. I., & Lewis, C. (2015). The development of social understanding. In L. Liben, U. Müller, & R. M. Lerner (Eds.), *Handbook of child psychology and developmental science: Cognitive processes: Vol. 2* (10th ed., pp. 381–424). Hoboken, NJ: John Wiley & Sons, Inc.

Case, R. (1992). The role of the frontal lobes in the regulation of cognitive development. *Brain and Cognition, 20*(1), 51–73. https://doi.org/10.1016/0278-2626(92)90061-P

de Villiers, J. G., & Pyers, J. E. (2002). Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief-understanding. *Cognitive Development, 17*(1), 1037–1060. https://doi.org/10.1016/S0885-2014(02)00073-4

Doebel, S. (2020). Rethinking executive function and its development. *Perspectives on Psychological Science, 15*(4), 942–956. https://doi.org/10.1177/1745691620904771

Fenici, M. (2015). A simple explanation of apparent early mindreading: Infants' sensitivity to goals and gaze direction. *Phenomenology and the Cognitive Sciences, 14*(3), 497–515.

Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review, 87*(6), 477. https://doi.org/10.1037/0033-295X.87.6.477

Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual perception: Further evidence for the Level 1–Level 2 distinction. *Developmental Psychology, 17*(1), 99. https://doi.org/10.1037/0012-1649.17.1.99

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

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development, 10*(4), 483–527. https://doi.org/10.1016/0885-2014(95)90024-1

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3 1/2-7 years old on a Stroop-like day-night test. *Cognition, 53*(2), 129–153. https://doi.org/10.1016/0010-0277(94)90068-X

Göregenli, M. (1997). Individualist-collectivist tendencies in a Turkish sample. *Journal of Cross-Cultural Psychology, 28*(6), 787–794. https://doi.org/10.1177/0022022197286009

Guajardo, N. R., Parker, J., & Turley-Ames, K. (2009). Associations among false belief understanding, counterfactual reasoning, and executive function. *British Journal of Developmental Psychology, 27*(3), 681–702. https://doi.org/10.1348/026151008X357886

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences, 21*(6), 803–831. https://doi.org/10.1017/S0140525X98001769

Halford, G. S., Wilson, W. H., Andrews, G., & Phillips, S. (2014). *Categorizing cognition: Toward conceptual coherence in the foundations of psychology*. MIT Press.

Halford, G. S., Phillips, S., Wilson, W. H., McCredden, J., Andrews, G., Birney, D., … Bain, J. D. (2007). Relational processing is fundamental to the central executive and is limited to four variables. *The cognitive neuroscience of working memory* (pp. 261–280).

Harris, P. L. (2012). *Trusting what you're told: How children learn from others*. Harvard University Press.

Heyes, C. (2014). False belief in infancy: A fresh look. *Developmental Science, 17*(5), 647–659. https://doi.org/10.1111/desc.12148

Kagitcibasi, C. (1996). *Family and human development across cultures: A view from the other side*. Mahwah: Lawrence Erlbaum.

Kagitcibasi, C., & Ataca, B. (2005). Value of children and family change: A three-decade portrait from Turkey. *Applied Psychology, 54*(3), 317–337. https://doi.org/10.1111/j.1464-0597.2005.00213.x

Lee, K. (2013). Little liars: Development of verbal deception in children. *Child Development Perspectives, 7*(2), 91–96. https://doi.org/10.1111/cdep.12023

Mischel, W., Shoda, Y., & Rodriguez, M. I. (1989). Delay of gratification in children. *Science, 244*(4907), 933–938. https://doi.org/10.1126/science.2658056

Morton, J. B., Trehub, S. E., & Zelazo, P. D. (2003). Sources of inflexibility in 6-year-olds' understanding of emotion in speech. *Child Development, 74*(6), 1857–1868. https://doi.org/10.1046/j.1467-8624.2003.00642.x

Muggeo, V. M. (2008). Segmented: An R package to fit regression models with broken-line relationships. *R News, 8*(1), 20–25.

Müller, U., Liebermann-Finestone, D. P., Carpendale, J. I., Hammond, S. I., & Bibok, M. B. (2012). Knowing minds, controlling actions: The developmental relations between theory of mind and executive function from 2 to 4 years of age. *Journal of Experimental Child Psychology, 111*(2), 331–348. https://doi.org/10.1016/j.jecp.2011.08.014

Muller, U., & Kerns, K. (2015). The development of executive function. In L. S. Liben, U. Muller, & R. M. Lerner (Eds.), *Handbook of child psychology and developmental science, Vol. 2: Cognitive processes* (7th ed., pp. 571–623). Hoboken, NJ: Wiley.

Nelson, K., & Fivush, R. (2004). The emergence of autobiographical memory: A social cultural developmental theory. *Psychological Review, 111*(2), 486. https://doi.org/10.1037/0033295X.111.2.486

Pascual-Leone, J. (1970). A mathematical model for the transition rule in Piaget's developmental stages. *Acta Psychologica, 32*, 301–345. https://doi.org/10.1016/0001-6918(70)90108-3

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

Piaget, J., & Inhelder, B. (1971). *Mental imagery in the child: A study of the development of imaginal representation*. Routledge.

Povinelli, D. J., & Vonk, J. (2003). Chimpanzee minds: Suspiciously human? *Trends in Cognitive Sciences, 7*(4), 157–160. https://doi.org/10.1016/S1364-6613(03)00053-6

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72*(3), 655–684. https://doi.org/10.1111/1467-8624.00304

Woolley, J. D., & Ghossainy, M. E. (2013). Revisiting the fantasy–reality distinction: Children as naïve skeptics. *Child Development, 84*(5), 1496–1510. https://doi.org/10.1111/cdev.12081

Zelazo, P. D. (2004). The development of conscious control in childhood. *Trends in Cognitive Sciences, 8*(1), 12–17. https://doi.org/10.1016/j.tics.2003.11.001

Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols, 1*(1), 297. https://doi.org/10.1038/nprot.2006.46

Zelazo, P. D., & Müller, U. (2011). Executive function in typical and atypical development. *The Wiley-Blackwell handbook of childhood cognitive development* (2nd ed., pp. 574–603). West Sussex: Wiley-Blackwell.

Zelazo, P. D., Müller, U., Frye, D., & Marcovitch, S. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development*, 1–137. https://doi.org/10.1111/j.15405834.2003.06803003.x