Martin-Ordas G (2018) "First, I will get the marbles." Children's foresight abilities in a modified spoon task, *Cognitive Development*, 45, pp. 152-161.

DOI: 10.1016/j.cogdev.2017.07.001

"First, I will Get the Marbles."
Children's Foresight Abilities in a Modified Spoon Task

Gema Martin-Ordas

Newcastle University

University of Stirling

Author note

Gema Martin-Ordas, Centre for Behaviour and Evolution, Newcastle University, and Department of Psychology, University of Stirling, UK. This work was supported by a Marie Sklodowska-Curie research fellowship.

Correspondence concerning this article should be addressed to Gema Martin-Ordas, Centre for Behaviour and Evolution, Institute of Neuroscience, Henry Welcome Building for Neuroecology, Framlington Place Newcastle University, NE2 4HH, United Kingdom.

Contact: gema.martin-ordas@ncl.ac.uk; gema.martin-ordas@stir.ac.uk

Word count (including references): 7942

Accepted for publication in *Cognitive Development* published by Elsevier. The final published version is available at: https://doi.org/10.1016/j.cogdev.2017.07.001

ABSTRACT

Previous methodologies used to investigate future thinking (i.e., one-step "spoon test") do not directly assess temporal reasoning. Consequently, the extent to which foresight is required to solve these tasks has been questioned. In the current study, 3-, 4- and 5-year-olds were presented with a *two-step* "spoon test": to secure a future need (e.g., play with a marble run game), children *first* had to obtain a key that allowed them *next* to access the marbles. By the age of 4 children selected the key; however, it is only by the age of 5 that children reasoned about the temporal sequence of future events *and* selected the key. Temporal reasoning, memory for the past events and age significantly contributed to predict children's ability to select the correct item. These findings suggest that temporal reasoning is crucial to assess future thinking and that item-choice measures alone might not involve foresight.

Keywords: spoon test, temporal reasoning, planning, memory, preschoolers

1. INTRODUCTION

The ability to envision the future and plan for future needs is a critical cognitive skill for our daily life (Suddendorf & Corballis, 2007). For example, thinking about what I will cook for dinner tonight allows me to plan what I will need to buy at the supermarket. Developmental research has shown that young children can think and plan for future needs (e.g., Atance, 2015, for a review). One of the most commonly used paradigms to assess future thinking in children is illustrated by Tulving's (2005) "spoon test." This test is based on the following scenario: A young girl dreams that she is at a party where all the guests are being served a delicious chocolate pudding. However, to eat the pudding, the guests must have their own spoon and the young girl does not have one. That night, she falls asleep while holding a spoon in her hand because she wants to avoid making the same mistake again. Bringing the spoon represents an instance of future thinking (i.e., foresight) because it implies *envisioning* a need that will occur in the *future*.

Based on the "spoon test" paradigm, researchers have shown that between ages 3 and 5 there is an important increase in children's abilities to plan for a future event (e.g., Atance & Meltzoff, 2005; Atance, Louw & Clayton, 2015; Busby & Suddendorf, 2005; Russell, Alexis, & Clayton, 2010; Scarf, Gross, Colombo, & Hayne, 2013; Suddendorf, Nielsen & von Gehlen, 2011). For example, Suddendorf and colleagues (2011) presented children with a locked box that had a triangular keyhole. The experimenter showed children that by using a red triangle key the box could be opened and then, they could obtain a sticker previously placed inside the box. A future need for a key was created by the experimenter pretending to have broken the key that had just been used. Then, children were taken to a second room. After 15 min they were told that they were going to go back to the first room and they were asked to select one of four keys to take with them. The correct object was a yellow triangle key, and the incorrect objects were a red cross key, a red square key, and a yellow square key. Suddendorf et al.’s (2011) results showed that whereas 4-year-old children could successfully select the correct key to solve the future problem (e.g., unlock the box to obtain the sticker), 3-year-olds failed to do so.

One concern with these “spoon test” methodologies—hereafter referred to as one-step “spoon test” tasks- is that it is unclear the extent to which children need to think about a *future event* in order to make a successful choice. In these tasks, selecting the correct item may only indicate that children know that, for example, the key is useful for unlocking the box *now* without having to represent its use in a future event (McCormack & Hoerl, 2011; Hudson, Mayhew & Prabhakar, 2011). This is because the correct choice for the future is the same as for the present moment (Redshaw & Suddendorf, 2013). In this regard, McCormack and Hoerl (1999; McCormack & Hoerl, 2011) have argued that assessing a concept of *future* involves setting up situations in which understanding *before-and-after relationships* between different events in time is required—henceforth “temporal reasoning.” Being able to appreciate that events occupy particular temporal locations in novel sequences is what allows us to foresee the *future* and is crucially important to how we make decisions about the future (McCormack & Hoerl, 2011). Returning to the earlier example, I know that right now I do not have money in my wallet; therefore, *first* I should go to the cash machine and *next* I should go to the supermarket. Failing to imagine and organize these two events in the correct temporal sequence will put me in an uncomfortable situation at the supermarket.

When does the ability to reason about temporal sequences start to develop? Research has shown that children remember temporal information from a very early age (e.g., Bauer & Mandler, 1989; Benson, 1997; Fivush 1984; Friedman, 1977; Nelson & Gruendel, 1986). For example, in an elicited-imitation of event sequences paradigm, Bauer and Mandler (1989) showed that children younger than 2 years of age recalled the order in which particular sequences of actions (e.g., making a picture) occurred. In contrast, children's ability to *use* the recalled temporal information in past-oriented decision-making tasks has been reported to emerge only by age 5 (e.g., Hoerl & McCormarck 2011; McCormack & Hoerl, 2005, 2007; McColgan & McCormack, 2008; Povinelli 2001). For example, McCormack and Hoerl (2007) presented 4- and 5-year olds with an object-location task. Children were introduced to two dolls (e.g., John and Peter) that always performed actions (e.g., brush their hair) in a specific order (e.g., John first, Peter last). Crucially, in this task being able to correctly identify the location of the object (e.g., brush) depended on the recollection of the order in which the two dolls acted. McCormack and Hoerl (2007) found that whereas 5-year-olds successfully identified the location of the object, 4-year-olds failed to do so.

A similar developmental pattern has been found in relation to the use of temporal information in future-oriented decision-making tasks (e.g., Klahr & Robinson, 1981; McColgan & McCormack, 2008; McCormack & Atance, 2011; McCormack & Hanley, 2011; Scholnick, Friedman, & Wallner-Allen, 1997; Welsh, 1991). For example, Welsh (1991) presented 3 to 12 year olds with the Tower of Hanoi problem. In this task, children were shown a wooden structure with 3 pegs and were asked to transfer the disks from one peg to another peg in order to achieve a goal state (i.e., the same disk pattern as in the experimenter's wooden structure). Importantly, children could only move one disk at a time and a disk could never be placed on top of a smaller one. Their results showed that by the age of 4 children could reason about the order of the required steps to solve the Tower of Hanoi problem—at least, when the minimum number of moves was necessary to succeed (Welsh, 1991; see also Carlson, Moses & Claxton, 2004).

Using a very different paradigm, McColgan and McCormack (2008) assessed preschoolers' ability to reason about temporal order in a route-planning task. In this task, 3-, 4- and 5-year-olds were presented with a doll character that was planning to visit a toy zoo and wanted to take a picture of one of the animals (e.g., kangaroo). Importantly, the doll did not have a bag to carry the camera. Thus, children were asked to identify a location where the doll could leave the camera so she could pick it up before getting to the kangaroo location. McColgan and McCormack's (2008) findings indicated that whereas 5-year-olds identified the correct locations to leave the camera, 3- and 4-year-olds failed to do so (see also McCormack & Hanley, 2011). Using a similar paradigm, Prabhakar and Hudson (2014) have recently shown that 4-year-olds succeeded at temporally organizing two goals (e.g., deciding the order in which to visit two locations in order to buy someone a present) under high working memory demands and when the contingencies between the goals were not explicitly highlighted. Three year olds did so only when the working memory demands were low and the contingency between the events was explicit (Prabhakar & Hudson, 2014). These results suggest that the cognitive demands involved in a temporal reasoning task can dramatically affect children's performance.

Altogether the above-mentioned studies show that the ability to reason about before-and-after relationships in tasks that involve achieving *immediate* goals improves between ages 3 to 5. Importantly, what these studies do not directly address is whether children can reason about temporal sequences in response to a problem that is not currently manifest—a crucial feature of future thinking abilities (e.g., Suddendorf & Corballis, 2007). Thus, the goal of the current study was to assess preschoolers' ability to plan for a future event in a task that required reasoning about a temporal sequence. To do so children were presented with a *two-step* "spoon test" paradigm that involved visiting two rooms (e.g., Suddendorf et al., 2011). To achieve a final goal (e.g., play with a marble run game), children *first* had to obtain a "tool" (e.g., key) so they could *next* obtain a second "tool" (e.g., marbles). The motivation to use the "spoon test" paradigm was based on previous work on future thinking in children (e.g., Atance & Meltzoff, 2005; Atance, Louw, & Clayton, 2015; Busby & Suddendorf, 2005; Russell, Alexis, & Clayton, 2010; Scarf, Gross, Colombo, & Hayne, 2013; Suddendorf, Nielsen & von Gehlen, 2011). These studies assessed planning as the selection of a correct item for a future use. In addition to (1) planning (i.e., item-choice), the two-step "spoon test" paradigm allowed the investigation of the two other aspects of future thinking abilities: (2) temporal reasoning (i.e., order in which to visit the rooms) and (3) memory (i.e., recollection for the contents of the rooms).

The general procedure consisted of showing children a locked box with marbles and a bag with crayons in a first room (e.g., Moon room). Next, in a second room (e.g., Rainbow room), they were presented with a marble run game. Note that only those children preferring the marble run game over the coloring game were included in the study. The rationale for this methodological decision was to ensure that children preferred the task that involved temporal reasoning (i.e., marble run game). After visiting both rooms, children were taken to an open area where they could not see either room. Children were told that they would be going back to both rooms after the completion of a 15 min sand-timer's cycle (e.g., Suddendorf et al., 2011; Redshaw & Suddendorf, 2013). Children were also provided with a bag and they were told that they would have to take it with them back to the rooms. Despite not being a main measure in the current study, this opportunistic setup was used to examine whether children would *spontaneously* remember to bring the bag after the 15-min delay. Previous research has shown that the ability to remember to perform an action in the future—prospective memory (e.g., Kerns, 2000)- improves between ages 3 and 5 (e.g., Mahy, Moses & Kliegel, 2014). However, the extent to which prospective memory is necessary for future thinking is still an open question. For example, Nigro, Bradimonte, Cicogna and Cosenza (2014) showed a positive relationship between both abilities; however, Atance and Jackson (2009) showed no relation.

Next, children were assigned to one of the 3 following conditions: Pre-delay decision, Post-delay decision and Immediate decision. In the Pre-delay decision condition children were asked *before* the 15-min delay to choose one of 3 items (e.g., a key, a coloring paper, a block) to put inside the bag. In the Post-delay decision condition children were asked to do so *after* the 15-min delay. The motivation for this manipulation was to assess the effect of the time delay on children's performance. The memory load when the decision is made before the delay should be smaller compared to when the decision is made after the delay. Consequently, one might expect the former condition to rely more on planning than memory abilities, and the latter to rely on both abilities. Once children selected one of the items, they were asked two sets of questions: the temporal reasoning questions (e.g., *"Which room would you like to go to first?", "Which room would you like to go to next?"*) and the memory questions (e.g., *"Do you remember what we saw inside the Rainbow room?", "Do you remember what we saw inside the Moon room?"*). The rationale for always asking these questions after children selected an item was to avoid the children having the problem being cued by the memory questions. Finally, in the Immediate decision, there was no spatial (i.e., games and items were presented in one room) or temporal displacements (i.e., there was no 15 min delay). The motivation for having this control condition was to assess that children could solve the problem when the cognitive demands (e.g., memory) were low.

Age-related differences in children's performance in planning, temporal reasoning and memory were predicted. Of particular interest was when the *combined* ability to successfully select the correct item *and* correctly reason about sequences of future events emerges in development; that is, whether it parallels the development of planning and/or temporal reasoning. This is a crucial aspect of the present task. As mentioned before, previous one-step spoon tasks do not necessarily assess whether participants think of the events as being *future* events (McCormack & Hoerl, 2011). Thus, assessing preschoolers' performance when planning entails temporal reasoning would help to shed light on this issue.

The contribution of memory and temporal reasoning abilities to performance in the planning component was also analyzed. Previous research has suggested that memory rather than future-thinking ability drives performance in one-step "spoon test" tasks (e.g., Atance & Sommerville, 2014). Therefore, if as McCormack and Hoerl (1999, 2011) have argued, being able to reason about temporal order plays a crucial role in foreseeing the future, item choice in the current task should also draw on future thinking abilities. This line of reasoning follows because a successful foresight response would involve not only choosing the correct item but also deciding the correct order in which to visit the rooms. Consequently, it was hypothesized that age-related differences in planning would be not only be due to memory abilities but also to temporal reasoning. In contrast, if children's performance in two-step "spoon test" tasks is mainly driven by memory—like in one-step "spoon test" tasks- age-related differences in planning would be due to memory rather than temporal reasoning.

2. METHODS

2.1. Participants

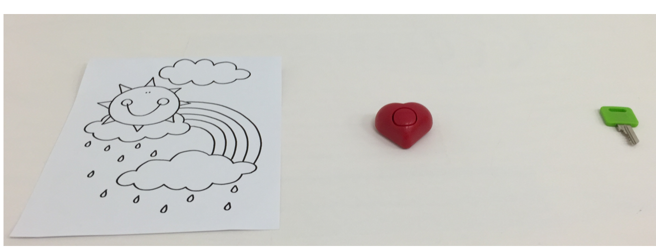
One hundred and two typically developing children were recruited, with two excluded due to preferring playing with crayons and paper over playing with a marble run game, resulting in a final sample of 100 children (41 females; 59 males) aged 3 (M=41months, SD=3.1months, n=36), 4 (M=52, SD=3.9, n=36) and 5 (M=67, SD=3.3 n=28). The ethnic composition of the sample was 93% Caucasian and 7% Asian and all participants were predominantly middle class, and fluent in English. Children were tested individually in the child lab facilities at the Institute of Neuroscience. The experiment received ethical approval from the Newcastle University's Faculty of Medical Sciences Ethics Committee. Parents provided written informed consent for their children's participation, and children also provided their verbal assent.

2.2. Materials and procedure

A marble run game, 6 marbles, a transparent box (7x4x4cm) with a lock, a key, 6 crayons, 1 coloring page, a block (3x3x1.5cm) and a red plastic bag were used in the present study (see Figure 1). A 15 min purple sand-timer (16cm x 8.3cm x 7.3cm) was used to indicate *when* children and Experimenter (E) would return to the rooms. The experiment took place in 4 different rooms/areas: Room 1, Room 2, Room 3 and Area 1. Room 1, Room 2 and Area 1 were located on the second floor of the Institute of Neuroscience and Room 3 was on the first floor. Participants first visited Room 1 and then Room 2. Next, participants went to Area 1 and, then, waited 15 min in Room 3 before returning to Rooms 1 and 2.

(a)

(b)



(c)

Figure 1. Box with marbles and bag with crayons presented in Room 1 (a), marble run game presented in Room 2 (b) and coloring page, block and key presented in Area 1 (c)

One of two pictures— either one depicting a moon or one depicting a rainbow- was placed on the door of Room 1 and the other picture was placed on the door of Room 2. For 50% of the children, Room 1 had the picture of the moon and Room 2 had the picture of the rainbow, and for the other 50% the order was reversed. E showed children both rooms "*Look, there are two rooms: the Rainbow room and Moon room and we are going to get to play inside these rooms. Would you like to go inside the Rainbow/Moon room and see what is inside? Let's go inside the Rainbow/Moon room!*" Then, E and child entered Room 1. First, children's preference for the two games was established by showing them a picture with crayons and a coloring page and a picture with a marble run and marbles, and asking them "*What do you like best: playing with crayons and papers or playing with marbles and marble runs?*" Children could answer verbally or point to their preferred game. Having this information ensured that children's item choices were based on their preferences. Note that it was crucial for the design that children preferred the marble run game over the coloring game. This is because a successful performance required temporal reasoning for the former but not necessarily for the latter. Therefore, those participants who chose the coloring game were not included in the experiment.

Next, E said *"Look, here is a box with some marbles. But we can't get them out of the box right now because the box is locked! We may need something to open it! Look, here is a bag with some crayons! But we can't use them right now! We may need something to draw on!"* The order in which the E showed children the marbles and the crayons was randomized. Next, E said *"I just remembered that I need to do some work in this room. I will be done really soon"* and for 1 min E pretended that she was filling out a form. The rationale for having this waiting time was to give children the possibility to manipulate the locked box containing the marbles and the bag with the crayons. This would allow them to realize that they could not play with either of them *right now* because something else was needed—a key and a coloring paper. After 1 min E said "*OK, I am done working here. Would you like to come with me to the Moon/Rainbow room?*" In Room 2, E said: *"Look, a marble run! But we can't play with it right now because there are no marbles! Oh I just remembered that I still need to do some work. I will be done really soon."* Then, E pretended to fill out a form for 1 min. As before, the rationale for having this waiting time was to give children the opportunity to explore the marble run, so they understood that there were no marbles and they could not play the game *right now*. When 1 min was up, E said: *"OK. I am done working here. Let's go outside."* Each child was then randomly assigned to one of three conditions defined by *when* children chose an item to take into the rooms:

1. *Pre-delay decision condition*: E and child went to Area 1—note that from this area participants did not have visual access to Room 1 or Room 2. Then, E showed a sand-timer to the child *"Look at this! It is a sand-timer! I am going to flip it so you can see how the sand falls in the bottom part!"* Next E explained *"when all the sand falls in the bottom part, we are going to go back to the two rooms we just visited and we are going to take this bag into the rooms with us!"* E showed the child a small red bag and the instructions were repeated. Next, E took a key, a coloring paper and a block out of a drawer and asked the following 3 sets of questions:

(a) Planning (i.e., item choice) question: E asked *"You can choose one of these three things and put it in the bag. Which one are you going to choose?"* One children made their choice, they put the item in the bag.

(b) Temporal reasoning questions: E asked *"Which room would you like to go to first?"* and *"Which room would you like to go to next?"* At no point did E mention the items/problem that children faced in both rooms.

(c) Memory questions: E asked *"Do you remember what we saw inside the Rainbow room?", "Do you remember what we saw inside the Moon room?"* The order in which the memory questions were asked was randomized across subjects. Note that the planning question, the temporal reasoning questions and the memory questions were always asked in this fixed order. This was done to avoid the children having the problem being cued by the memory questions.

Once children answered the questions, E explained *"I just remembered that I have to make a phone call. It won't take long but maybe you and your mum/dad can play with some games while I make the phone call. You should take the bag and sand-timer with you and keep on eye on the timer because, remember, when all the sand falls in the bottom part we will be returning to these two rooms and we will be talking the bag with us."* E took the child and parent to Room 3 where they played for 15 min with unrelated games. Parents were instructed not to mention anything about the experimental task while playing with the child.

After the 15 min, E went back to Room 3 and proceeded with the prospective memory aspect of the task by asking *"Are you ready to come upstairs with me?"* If the child did not spontaneously remember to grab the bag, E said *"Is there anything you need to bring upstairs with you?"* Only 14% of the children spontaneously remembered to take the bag and, as such, this low performance might indicate that remembering to take the bag is not a good measure of prospective memory. Thus, in order to avoid spurious effects of the conditions manipulated in the current study, this measure was not further analyzed. Next, E and child went into Room 1 and Room 2 in the order in which the child had previously decided to do so. In those cases in which children chose the correct item but indicated the incorrect order, E asked "*what do we do now?*" Finally, in those cases in which children chose the paper or the block, E explained to the child which item s/he needed to unlock the box and in which order they should visit the rooms.

(2) *Post-delay decision condition:* The same procedure as for the Pre-delay decision condition was used—with the *only* difference that in the Post-delay decision condition E asked the 3 sets of questions (i.e., planning, temporal reasoning and memory) *after* the 15 min delay. As in the Pre-delay decision condition, the order in which these questions were asked was fixed.

(3) *Immediate decision condition*: In this control condition children were presented with all the elements of both games in the same room. The procedure was the same as for the Pre-delay decision and Post-delay decision conditions. However, there was no spatial displacement nor temporal delay between showing children the games—marble run, locked box with marbles, bag with crayons-, presenting them with the three items—key, coloring paper, and block- and asking them the planning question. In this condition, the memory questions were not asked. The rationale behind the Immediate decision condition was to assess whether children could solve the problem (1) when all

the elements of both games were presented in the same room and (2) when no delay was involved between showing them the games and letting them choose one of the three items.

2.3. Scoring and analyses

Sessions were video-recorded. The data were scored according to three task variables, which reflected planning, temporal reasoning and memory. To assess planning abilities, children received a score of 1 if they chose the correct item in the Pre-delay decision, Post-delay decision and the Immediate decision conditions, and a score of 0 if they chose the incorrect item. The correct item was considered to be the key since all participants included in the analyses preferred playing with the marble run to coloring with the crayons. See Table 1 for the percentage of children choosing one of three possible items[1].

Children's responses to the temporal reasoning questions were given a score of 1 if they stated that they wanted to visit the marble room first and the marble run room second. Those responses in which children stated the reversed order received a score of 0. For the answers to the memory questions a participants' overall recollection was calculated – that is, successfully remembering the content of both rooms was scored as 1 and failing to remember the contents of the rooms was scored as 0.

Table 1. Percentage of children choosing each of the three objects grouped by age group and condition and condition.

[1] Neither age (χ^2=7.21, *df*=4, *p*=.123) nor condition (χ^2=6.25, *df*=4, *p*=.183) affected children's decisions to choose one of the 3 items.

Age	**Conditions**	**Key**	**Coloring paper**	**Block**
3YO	Immediate	84	8	8
	Pre-delay	58	9	33
	Post-delay	58	9	33
4YO	Immediate	92	8	0
	Pre-delay	67	8	25
	Post-delay	67	8	25
5YO	Immediate	84	0	16
	Pre-delay	84	8	8
	Post-delay	84	16	0

Pearson chi-square tests were used to compare children's planning, memory and temporal reasoning in the Pre-delay decision, Post-delay decision and Immediate decision conditions as well as to analyze the effect of age. A three-way cross-tab (Chi-square test) was used to analyze the interactions between age and condition. A binary logistic regression test was used to analyze the effect of age, temporal reasoning and memory on planning abilities. Finally, binomial tests were run to assess whether children were above chance at choosing one of the three items (chance=33%). Statistical tests were two-tailed, and results were considered significant if $p<.05$.

3. RESULTS

3.1. The three task variables

Planning. Condition had an effect on children's choices (χ^2=5.90, *df*=2, *p*=.049; *Cramer's V*=.24). Children tended to make more mistakes in the Pre-delay decision (χ^2=5.70, *df*=1, *p*=.01; φ=-.29) and Post-delay decision (χ^2=3.34, *df*=1, *p*=.06, φ=-.21) conditions than in the Immediate decision condition. However, no significant differences were found between the Pre-delay decision and Post-delay decision conditions (χ^2=.42, *df*=1, *p*=.516).

Age also had an effect on children's choices (χ^2=6.04, *df*=2, *p*=.049; *Cramer's V*=.24). However, post-hoc analyses revealed that only 5 year olds performed better than 3 years old (χ^2=6.04, *df*=1, *p*=.014; φ=-.31). The interaction between age and condition was not found to be significant (χ^2=3.22, *df*=4, *p*=.53). Yet all 3 age groups chose the key significantly above chance in the Immediate decision condition (Binomial test: *p*<.001 in all cases), 4- (*p*=.036) and 5- year olds (*p*=.024) did so in the Pre-delay decision condition, and only 5 year olds' performance was better than expected by chance in the Post-delay condition (*p*<.001) (see Figure 2).

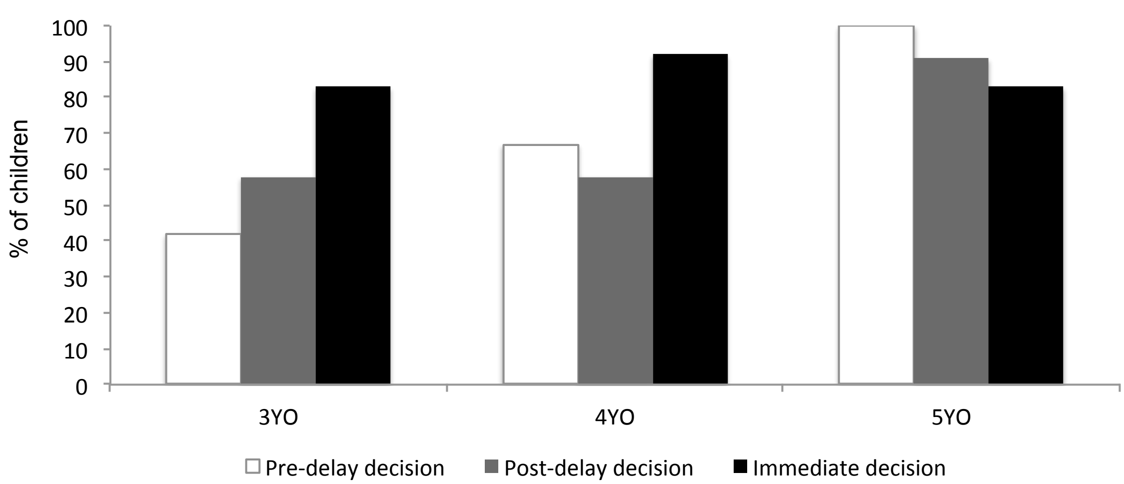


Figure 2. Percentage of 3-, 4- and 5-year-old children choosing the correct item in the Pre-delay decision, Post-delay decision and Immediate decision.

Temporal reasoning. Neither condition (χ^2=3.72, *df*=2, *p*=.156) nor age (χ^2=.165, *df*=2, *p*=.438) had a significant effect on children's responses to the temporal reasoning questions (see Figure 3). However, only 5 year olds indicated the correct order better than expected by chance (*p*=.019).

One aim of the present study was to investigate when the combined ability to choose the correct item and reason about sequences of future events emerges. To do so an overall score (i.e., successful performance in both planning and temporal reasoning) was calculated. Although age did not have a significant effect on children's responses (χ^2=4.22, *df*=2, *p*=.121), only 5 year olds performed better than chance (*p*=.052) (see Figure 3). The results also showed that condition affected children's responses (χ^2=7.07, *df*=2, *p*=.029; *Cramer's V*=.26). Children tended to make more mistakes in the Pre-delay decision (χ^2=7.44, *df*=1, *p*=.006; φ=-.33; 42% responded correctly) and Post-delay decision (χ^2=5.55, *df*=1, *p*=.018, φ=-.27; 48% responded correctly) conditions than in the Immediate decision (72% responded correctly). However, no significant differences were found between the Pre-delay decision and Post-delay decision conditions (χ^2=.26, *df*=1, *p*=.60). Finally, the interaction between age and condition was not found to be significant (χ^2=3.22, *df*=4, *p*=.52).

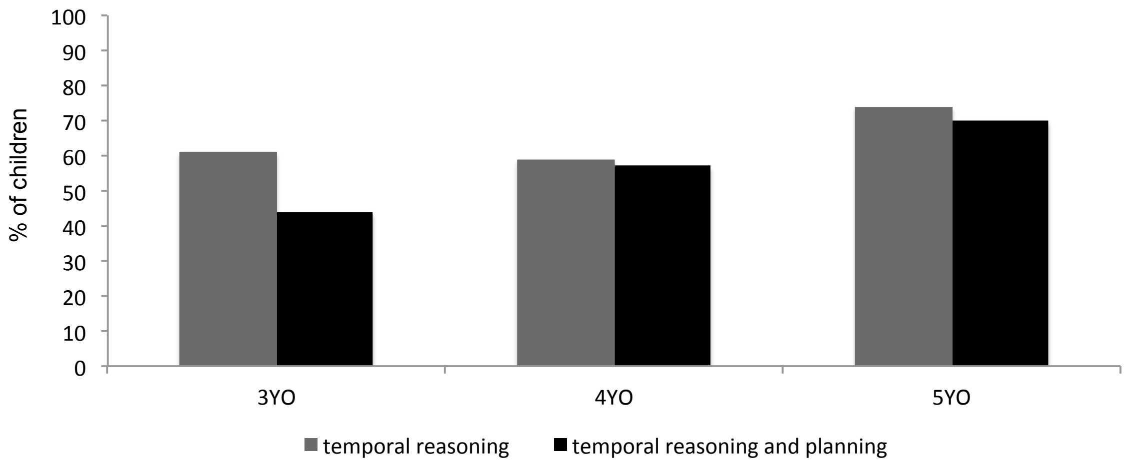


Figure 3. Percentage of children correctly responding to the temporal reasoning questions (grey bars) and to the combined ability to select the correct item and answer the temporal reasoning questions (black bars) grouped by age.

Memory. Condition did not affect children's recollection for the contents of the rooms (χ^2=.89, *df*=1, *p*=.343) but age did (χ^2=6.04, *df*=1, *p*=.049, φ=.31). Post-hoc analyses revealed that 4 year olds performed better than 3 year olds (χ^2=5.57, *df*=1, *p*=.018, φ=.34) and 5 year olds tended to perform better than 3 year olds (χ^2=2.88, *df*=1, *p*=.070) (see Figure 4). No differences were found between 4 year olds and 5 year olds (χ^2=.17, *df*=1, *p*=.74). In addition, only 4- (*p*<.001) and 5- year olds (*p*=.029) remembered the contents of the two rooms significantly above chance.

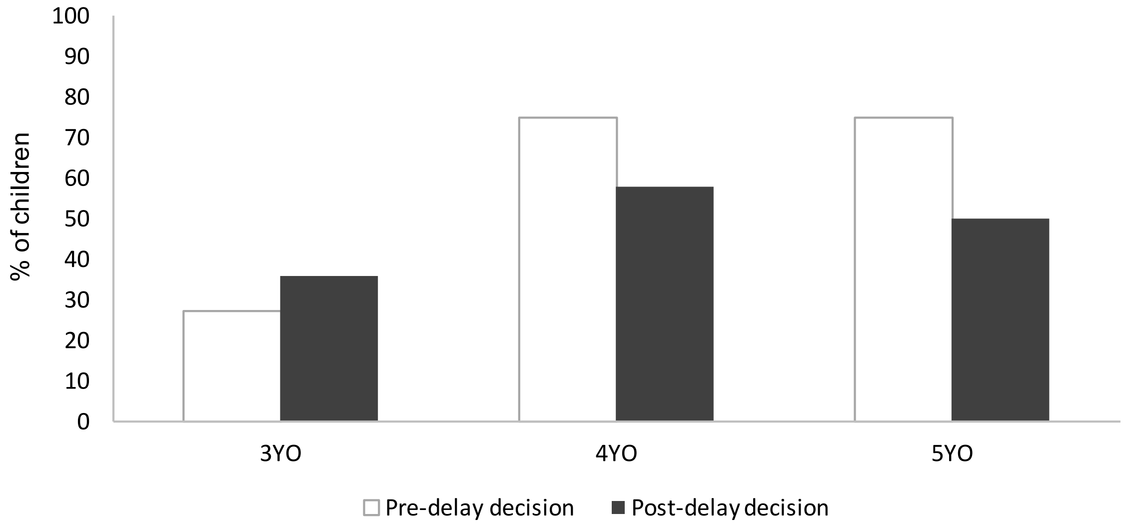


Figure 4. Percentage of children correctly responding to the memory questions in the Pre-delay decision and Post-delay decision conditions grouped by age.

3.2. The contribution of memory and temporal reasoning to planning

One of the goals in the current study was to investigate whether age-related changes in children's ability for planning were better explained by age-dependent changes in temporal reasoning abilities and/or memory abilities. A binary logistic regression analysis was run to predict children's responses to the item choice question (i.e., planning) from their age, the experimental condition and their responses to the temporal reasoning and memory questions. The regression model was significant (χ^2=19.01, *df*=4, *p*<.001). The model explained 36% (Nagelkerke R^2) of the variance in planning and correctly classified 75% of the cases. Importantly, the significant effect was driven by children's age (*Wald*=5.40, *SE*=.49, *p*=.020; *Odds ratio*: 3.13) and temporal reasoning (*Wald*=5.57, *SE*=.68, *p*=.018; *Odds ratio*: 5.01) and by a trend for children's memory (*Wald*=3.28, *SE*=.64, *p*=.070; *Odds ratio*: 3.24). Condition did not significantly predict children's planning responses (*Wald*=.022, *SE*=.67, *p*=.881; *Odds ratio*: .90). That is, children's ability to choose the correct item in the current task is predicted by age-related changes in planning and by their sequential reasoning and to some extent to memory abilities. The interactions between the 4 predictors were also calculated; however only age and temporal reasoning (*Wald*=3.64, *SE*=.20, *p*=.056; *Odds ratio*: 1.48) and age and memory (*Wald*=3.14, *SE*=.20, *p*=.076; *Odds ratio*: 1.43) tended to show an interaction effect.

4. DISCUSSION

In this study, planning (i.e., item choice), temporal reasoning (i.e., order in which to visit the rooms) and memory for the critical past information (i.e., content of the rooms) were assessed. All 3 age groups chose the correct item (i.e., key) when no spatial displacement or temporal delay was involved (i.e., Immediate decision). Four and 5 year olds chose the correct item when there was spatial displacement (i.e., Pre-delay decision) and only 5 year olds did so when there was both spatial displacement and temporal delay (i.e., Post-delay decision). Similar to Suddendorf et al. (2011), more children succeeded when there was no temporal delay between the presentation of the problem and the item-choice question (i.e., Immediate decision condition) compared to when there was a temporal delay between both events (i.e., Pre-delay decision and Post-delay decision conditions). However, no differences in performance were found between the Pre-delay and Post-delay decision conditions. Consistent with previous studies on temporal reasoning (e.g., McColgan & McComarck, 2008), only 5 year olds performed better than expected by chance at indicating the order in which to visit the rooms. Also in keeping with previous studies (e.g., Atance & Sommerville, 2014), children's memory abilities differed with age. Finally, children's capacity to choose the correct item in a two-step "spoon test" task was explained by age, temporal reasoning skills and memory.

Similar to one-step "spoon test" tasks (e.g., Suddendorf et al., 2011) by the age of 4 children use the information of past events to select an item for a future use. Importantly, 3 year olds did not fail the Pre-delay and Post-delay decision conditions because of poor skills to solve sequential reasoning problems—they successfully chose the correct item in the Immediate decision condition, that is, when the cognitive demands were not as high as in the Pre-delay and Post-delay decision conditions (see Prabhakar & Hudson, 2014 for similar findings). One possibility is that younger children changed their game preference during the delay and that after 15 min coloring was more appealing than playing with the marble run. If this were true, children would have chosen the coloring paper. A closer look at children's responses revealed that only 16% of the 3 year olds who chose one of the incorrect items chose the coloring paper—suggesting that a change in game preference was not underlying children's choices.

Another possible account is that poor memory abilities diminished 3 year olds' performance in the current task. In fact, their recollection for the content of the rooms tended to be poorer than older children's memories. Yet it is equally possible that 3 year olds failed the Pre-delay and Post-delay conditions because of their immature temporal reasoning abilities. As a matter of fact, they did not perform better than expected by chance at indicating the order in which to visit the rooms. Moreover, when the combined ability for planning and temporal reasoning was analyzed, 3 year olds still failed to perform better than chance—suggesting that when selecting the correct item they might not have done it while thinking about its future use. Strikingly, 4 year olds also failed the temporal reasoning questions in spite of choosing the correct item in the Pre-delay decision condition.

Yet one could argue that children did not have to reason about the order in which to visit the rooms, rather they simply had to remember the order in which they visited the rooms. Recall that children visited first the room with the marbles and next the room with the marble run. This explanation would imply a relation between children's responses to the memory question and children's responses to the temporal reasoning question. Crucially, a contingency coefficient analysis failed to show a significant relation between both measures (C=.011, df= 1, p=.799)—suggesting that the same mechanisms were not necessarily involved in both measures. Another possibility is that the order in which to visit the rooms was not a normative order because there are two possible "correct" responses to the temporal reasoning questions: (1) first unlock the marble box and then take the marbles to the marble run room; and (2) first take the marble run to the marble box room and then unlock the box. Thus, either response could indicate foresight. If children who chose the key had considered the second solution, then they should have answered Experimenter's question "*what do we do now?*" by stating that they wanted to take the marble run to the marble box room. However, 100% of the 3 year olds and 80% of the 4 year olds indicated that they could use the key to unlock the marble box and then bring the marbles into the marble run room. Thus, children assumed that the normative sequence was to carry the marbles to the marble run room rather than the marble run to the marble box room. Crucially, they were only able to indicate the correct order once they were inside the marble run room and realized that the key had no use in such room—suggesting that their item choices did not necessarily incorporate the temporal reasoning component. Likewise, this finding also opens up the possibility that item-choice measures alone might not reflect foresight—as suggested by McCormack and Hoerl (2011).

The findings discussed so far indicate that both memory and temporal reasoning contributed to children’s item choices. This observation was further confirmed when the different measures were pitted against each other. In keeping with previous studies (e.g., Atance & Sommerville, 2014), memory tended to predict children’s choices. These analyses also confirmed that in addition to memory, temporal reasoning contributed to explain children’s item-choice responses—supporting McCormack and Hoerl’s (2011) claim that future-thinking tasks should include a temporal reasoning component. In the same vein as Atance and Sommeville (2014), age was also analyzed; however contrary to their findings age did predict children’s item choices—suggesting that the development of planning abilities was also at play in the current task. Altogether these results raise the possibility that spoon tasks only including item-choice measures may not be sufficient to test foresight skills and that additional measures such as temporal reasoning have to be included; otherwise, and according to Atance and Somerville’s (2014) findings, it is unclear the extent to which other cognitive skills besides memory are being tested in one-step spoon tasks.

It is possible, however, that both planning and temporal reasoning represent two differentiable forms of foresight. As shown here, 4 year olds can successfully decide which item they will need in a future situation but fail to decide the order in which to carry out a sequence actions—suggesting that each ability follows different developmental patterns. Yet what is being addressed in the present study is the extent to which planning—as measured in item-choice tasks- entails a concept of future. In this regard, the combined planning and temporal reasoning score used here adds to the developmental literature of future thinking. Recall that only 5 year olds chose the key *and* correctly indicated the order in which to visit the rooms. Thus, if tasks requiring “a quite specific temporal sequencing ability” (McCormack & Hoerl, 2011, page 137) are testing the concept of future one would be tempted to suggest that the ability to plan for a future event may emerge later in preschool years than the ability to the select a correct item—as tested in one-step spoon tasks.

A possible limitation of the present study, though, is that the temporal reasoning and memory measures were embedded in the task and no other external measures of such skills were included—doing so would have helped not only to validate the measures within the task but also to investigate how these independent measures might contribute to children’s performance in the two-step spoon task. Crucially, the temporal reasoning and memory questions used here were the same as those used on previous studies addressing temporal reasoning and memory in preschoolers (e.g., McCormack & Hoerl, 2005; Martin-Ordas, Atance & Caza, 2017). Moreover, the developmental patterns reported here also replicate previous findings on temporal reasoning and memory (e.g., Atance & Sommerville, 2014; Martin-Ordas et al. in press; McColgan & McCormack 2008)—which offers support to the validity of the measures used in the present study. Yet future studies should address this issue by including both internal and external measures of memory and temporal reasoning within the spoon test tasks.

Finally, one would still need to explain why being asked to select an item *before* (Pre-delay decision) or *after* (Post-delay decision) the 15 min delay did not significantly affect children's responses to the planning and memory questions. Recall that the former was predicted to rely more on planning than memory and the latter on both. One tantalizing possibility is that both conditions *equally* relied on memory and planning. This is along the lines of previous studies investigating the contribution of the temporal distance into the future in children's item choices (Atance et al., 2015; Redshaw & Suddendorf, 2013). For example, Redshaw and Suddendorf (2013) argued that children rely on the same mechanisms when selecting an item that they can use in a near future (e.g., 5 min) and in a more distant future (e.g., 15 min). Another possibility is that both conditions *only* relied on memory. That is, once children have to draw on memories of past events in order to select an item for a future use, the length of the interval is not significant—at least, when the intervals are relatively short (e.g., Redshaw & Suddendorf, 2013; Scarf et al., 2013; Suddendorf et al. 2011). This latter account, although speculative, would be consistent with the idea that item-choices measures might only rely on memory and not on planning abilities. Notably, if either of these two explanations were to account for the present findings, one would have expected to see a significant interaction between age and condition for the planning and/or the memory questions; however, this is not what was found. This is, at least, an unexpected result given that age alone did affect children's planning and memory responses. Thus, at this point, the present results cannot offer a conclusive explanation.

In conclusion, children's ability to select an item that was useful for a future reward develops by the age of 4. However, children's combined ability to select the appropriate item *and* correctly reason about temporal sequences of future events only seems to emerge by the age of 5. Successful recollection of past events was not sufficient to explain age-related changes in children's item-choice responses—age and temporal reasoning abilities also contributed to children's performance. In this regard, the present study makes two important contributions that have implications for the understanding of preschoolers' future thinking skills. First, it is crucial to incorporate a temporal component to investigate future thinking. Second, alone item-choice measures might not involve foresight (see also Atance & Sommerville, 2014 for similar conclusions). Future research should extend the current findings by, for example, exploring in which other contexts besides tool-use tasks temporal reasoning plays a role in planning—this will also provide us with a more comprehensive picture of when in ontogeny this skill might develop.

Acknowledgments

I thank Cristina M. Atance and Quoc Vuong for insightful comments on earlier versions of this manuscript. Special thanks to all the children and parents who volunteered their participation.

REFERENCES

Atance, C. M. (2015). Young children's thinking about the future. *Child Developmental Perspectives*, *9*, 178-182.

Atance, C. M., & Jackson, L. K. (2009). The development and coherence of future-oriented behaviors during the preschool years. *Journal of Experimental Child*

Psychology, 102, 379–391.

Atance, C. M., & Meltzoff, A. N. (2005). My future self: Young children’s ability to anticipate and explain future states. *Cognitive Development*, *20*, 341–361.

Atance, C. M., & Sommerville, J. A. (2014). Assessing the role of memory in preschoolers’ performance on episodic foresight tasks. *Memory, 22*, 118–128.

Atance, C. M, Louw, A., & Clayton, N. S. (2015). Thinking ahead about where something is needed: New insights about episodic foresight in preschoolers. *Journal of Experimental Child Psychology, 129*, 98-109.

Bauer, P. J., & Mandler, J. M. (1989). One thing follows another: Effects of temporal structure on 1- to 2-year-olds’ recall of events. *Developmental Psychology, 25*, 197-206.

Benson, J. B. (1997). The development of planning: It’s about time. In S. L. Friedman & E. K. Scholnick (Eds.), *The developmental psychology of planning: Why, how, and when do we plan?* (pp 43–76). Mahwah, NJ: Erlbaum.

Busby, J., & Suddendorf, T. (2005). Recalling yesterday and predicting tomorrow. *Cognitive Development*, *20*, 362–372.

Carlson, S. M., Moses, L. J., & Claxton, L. J. (2004). Individual differences in executive functioning and theory of mind: An investigation of inhibitory control and planning ability. *Journal of Experimental Child Psychology, 87,* 299–319

Fivush, R., (1984). Learning about school: The development of kindergartners' school scripts. *Child Development, 55*, 1697-1709.

Friedman, W. J., (1977). Development of children’s understanding of cyclic aspects of time. *Child Development, 48,* 1593-1599.

Kerns, K., (2000). The CyberCruiser: An investigation of development of prospective memory in children. *Journal of the International Neuropsychological Society, 6*, 62–70.

Klahr, D., & Robinson, M. (1981). Formal assessment of problem-solving and planning processes in preschool children. *Cognitive Psychology, 13*, 113-148.

Hoerl, C., & McCormack, T. (2011). Time in cognitive development. In C. Callender (Ed.), *The Oxford handbook of time* (pp. 439–459). Oxford: Oxford University Press.

Hudson, J. A., Mayhew, E. M. Y., & Prabhakar, J. (2011). The development of episodic foresight: Emerging concepts and methods. *Advances in Child Development and Behavior, 40*, 95–137.

Mahy, C.E.V, Moses, L.J., & Kliegel, M. (2014). The development of prospective memory in children: An executive framework. *Developmental Review, 34*, 305-326.

Martin-Ordas, G., Atance, C.M., & Caza, J. (2017). Did the popsicle melt? Preschoolers' performance in an episodic-like memory task. *Memory,* http://dx.doi.org/10.1080/09658211.2017.1285940

McColgan, K. L., & McCormack, T. (2008). Searching and planning: Young children's reasoning about past and future event sequences. *Child Development, 79*, 1477–1497.

McCormack, T., & Atance, C. M. (2011). Planning in young children: A review and synthesis. *Developmental Review, 31*, 1-31.

McCormack, T., & Hanley, M. (2011). Children's reasoning about the temporal order of past and future events. *Cognitive Development, 26,* 299-314.

McCormack, T., & Hoerl, C. (1999). Memory and temporal perspective: The role of temporal frameworks in memory development. *Developmental Review, 19*, 154–182.

McCormack, T., & Hoerl, C. (2005). Children's reasoning about the causal significance of the temporal order of events. *Developmental Psychology, 41*, 54–63.

McCormack, T., & Hoerl, C. (2007). Young children's reasoning about the order of past events. *Journal of Experimental Child Psychology, 98*, 168–183.

McCormack T., & Hoerl, C. (2011). Tool use, planning, and future thinking in animals and children. In T. McCormack, C. Hoerl, & S. Butterfill (Eds.), *Tool use and causal cognition* (pp.129-147). Oxford, UK: Oxford University Press.

Nelson, K., & Gruendel, J. (1986). Children's scripts. In K. Nelson (Ed.), *Event knowledge: Structure and function in development* (pp. 21-46). Hillsdale, NJ: Erlbaum.

Nigro, G., Brandimonte, M., Cicogna, P.C., & Cosenza, M. (2014). Episodic future thinking as a predictor of children's prospective memory. *Journal of Experimental Child Psychology, 127*, 82–94.

Povinelli, D. J. (2001). The self: Elevated in consciousness and extended in time. In C. Moore & K. Lemmon (Eds.), *The self in time: developmental perspectives* (pp. 75-95). Mahwah, NJ: Erlbaum.

Redshaw, J., & Suddendorf, T. (2013). Foresight beyond the very next event: Four-year-olds can link past and deferred future episodes. *Frontiers in Psychology, 4*, 404.

Russell, J., Alexis, D., & Clayton, N.S. (2010). Episodic future thinking in 3- to 5-year-old children: The ability to think of what will be needed from a different point of view. *Cognition*, *114*, 56–71.

Scarf, D., Gross, J., Colombo, M., & Hayne, H. (2013). To have and to hold: Episodic memory in 3- and 4-year-old children. *Developmental Psychobiology, 55*, 125–132.

Scholnick, E. K., Friedman, S. L., & Wallner-Allen, K. E. (1997). What do they really measure? A comparative analysis of planning tasks. In S. L. Friedman & E. K. Scholnick (Eds.), *The developmental psychology of planning: Why, how, and when do we plan?* (pp. 127-156). Mahwah, NJ: Erlbaum.

Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: what is mental time travel and is it unique to humans? *Behavioral and Brain Sciences, 30*, 299-313.

Suddendorf, T., Nielsen, M., & von Gehlen, R. (2011). Children's capacity to remember a novel problem and to secure its future solution. *Developmental Science*, *14*, 26–33.

Tulving, E. (2005). Episodic memory and autonoesis: Uniquely human? In H. S. Terrace & J. Metcalfe (Eds.), *The missing link in cognition: Origins of self-reflective consciousness* (pp. 3–56). New York, NY: Oxford University Press.

Welsh, M. C. (1991). Rule-guided behavior and self-monitoring on the Tower of Hanoi disk-transfer task. *Cognitive Development, 6*, 59–76.