More Later: Delay of Gratification and Thought About the Future in Children


Published in:
Child Development

Document Version:
Publisher's PDF, also known as Version of record

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Publisher rights
Copyright 2021 the authors. This is an open access article published under a Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution and reproduction in any medium, provided the author and source are cited.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Download date: 12. Mar. 2021
More Later: Delay of Gratification and Thought About the Future in Children

Patrick Burns and Patrick A. O’Connor
Queen’s University Belfast

Cristina Atance
University of Ottawa

Teresa McCormack
Queen’s University Belfast

We investigated whether individual differences in future time perception and the detail with which future events are imagined are related to children’s delay of gratification. We administered a delay choice task (real rewards), a delay discounting task (hypothetical rewards), a novel future time perception measure, an episodic future thinking (EFT) interview and IQ measures to a sample of 7- to 11-year-olds (N = 132) drawn from a urban predominately white population in N. Ireland. We found a strong correlation between delay choice and delay discounting. Future time perception and EFT were related to delay discounting, however only the relation with future time perception survived controlling for age and IQ. Children who showed greater compression of future time periods were the steepest discounters.

People often face decisions that require striking a balance between an immediate reward and delayed consequences, such as whether to put this month’s pay bonus into a pension fund or spend it on a shopping trip. Decisions of this sort are called intertemporal choices and research has shown that broadly speaking, when faced with such decisions, many people tend to sharply discount future rewards in favor of immediate gratification (Frederick, Loewenstein, & O’Donoghue, 2002). Moreover, the longer the delay to the future reward the stronger the preference is for immediate gratification (Green, Fry, & Myerson, 1994).

A difficulty in delaying gratification is especially pronounced among children. The classic marshmallow task is the paradigmatic example of this difficulty in preschoolers (Mischel, Shoda, & Rodriguez, 1989), but there is now a wide-ranging literature with a variety of tasks attesting to the difficulties children, even beyond the preschool years, have in delaying gratification (e.g., Marchetti, Castelli, Santivo, & Massaro, 2014; Staubitz, Lloyd, & Reed, 2018). What then might explain the difficulties children have in making prudent intertemporal choices? In this study, we investigate the role of two aspects of thought about the future that have recently been linked to adult performance on intertemporal choice tasks. The first is future time perception, that is, how far away future time points feel subjectively (Kim & Zauberman, 2009, 2019). The second is episodic future thinking (hereafter EFT), the ability to imagine events that one may personally experience in the future (Tulving, 2005). Below, we outline the case for hypothesizing that individual differences in future time perception and EFT skills will be associated with children’s intertemporal choice, before going on to describe the details of this study. However, in order to explain our measures, it is necessary to begin by giving a description of intertemporal choice tasks.

Measuring Intertemporal Choice in Children

The defining feature of intertemporal choice tasks is their structure: individuals choose between a small sooner reward (SSR) and a large later reward (LLR). Studies of delay of gratification with
young children use one of two task types, the delay maintenance task (Mischel, 1966) or the delay choice task (Mischel & Metzner, 1962). In delay maintenance tasks, such as the marshmallow task, an initial preference for the delayed option is assumed and subsequent self-control is measured in the form of waiting times for the delayed option (participants can switch from their choice at any point). In the delay choice task, participants make a choice between the SSR and the LLR and this choice is binding; the measure of interest is how often the LLR is chosen. Somewhat surprisingly, the delay discounting task and the delay choice task show little to no association with one another (Duckworth & Kern, 2011; Toner, Holstein, & Heterington, 1977). In this study, we focus on children’s delay of gratification as measured by the delay choice task, along with its homolog in the adolescent and adult literature, the delay discounting task.

In delay discounting tasks typically used by adolescents and adults, participants make multiple choices involving various time delays and reward magnitudes (e.g., $10 now or $100 in a year). The simpler delay choice tasks that have been used by children have a similar choice structure but there are nevertheless many differences between these and the more complex delay discounting tasks, due to the constraints that apply when testing young children. In the delay choice tasks typically used by children, delays are usually in the order of minutes to hours and the rewards are always realized (e.g., stickers). Trial numbers are low and the dependent variable is typically the proportion of delayed options chosen.

Delay discounting tasks by contrast usually employ hypothetical monetary rewards, though the discounting of many other types of commodities has been studied (see Urminsky & Zauberman, 2015, for a review). The use of hypothetical rewards allows for the inclusion of large reward values, long delay periods and many more trials than is feasible in delay choice tasks. This in turn leads to different forms of dependent variable over and above the proportion of delayed options chosen: with enough trials at a given delay it is possible to calculate an indifference point for that delay, the value of the SSR such that the participant has no preference between that and the LLR. Indifference points can be interpreted as the present subjective value of an LLR (e.g., if participants are indifferent between $10 now and $100 in a year’s time, this suggests that the $100 in a year’s time currently has a subjective value of $10). If indifference points are calculated for a number of delays, then a mathematical function can be fitted that describes how the LLR is discounted (i.e., reduced in value) across time. Figure 1 displays three different mathematical functions commonly cited in the discounting literature: The hyperbolic and beta-delta functions have been gained prominence due to their putative descriptive fit, whereas the exponential discount function is regarded as a normative model of discounting. In Figure 1 the functions have been fitted to simulated indifference points from a sample discounting study.

Although it is beyond the scope of this article to evaluate the applicability of these different functions in a developmental context (see Burns, Fay, et al., 2019, for a detailed discussion), fortunately we can take an unbiased approach by fitting each participant’s data with any number of models and then perform Bayesian model selection (Schwarz, 1978) to select the best fitting model. Once the best fitting model for the data has been determined we can measure the area under the function curve as a summary statistic of the overall level of discounting. We refer to this as the model-based area under the curve (hereafter model-based AUC; see Gilroy & Hantula, 2018, for further details) to distinguish it from the more common method of measuring AUC using trapezoids, which we will refer to as point-based AUC (Myerson, Green, & Warusawitharana, 2001). As can be seen from Figure 1, if the LLR quickly depreciates in present subjective value the area subtended by the functions will be smaller

**Figure 1.** The points display simulated indifference points. Each of the three lines represents a class of discounting function that has been fitted to the indifference points using the nonlinear least squares method. Model selection proceeds by calculating Bayesian information criteria (BIC) scores for each function. In this hypothetical example, the beta-delta model provides the best fit (i.e., lowest BIC score). The area under the beta-delta curve is shaded and it can be calculated precisely using integral calculus to give a summary statistic of overall discounting.
(i.e., the functions will decline more steeply) than if the LLR had a relatively high present subjective value. Thus, AUC is a measure of the extent to which participants show impulsive choice (smaller AUC = more impulsive). Model-based AUC has several advantages over alternative measures of discounting such as point-based AUC or the log of the hyperbolic discount function parameter. Unlike point-based AUC, model-based AUC does not require the connection between indifference points to be interpolated, rather it takes an unbiased approach using statistical criteria to determine the shape of the discounting curve. The model-based AUC approach also avoids a common problem with the point-based approach whereby small differences in indifference points at longer delays have a disproportionate influence on the final measure (Borges, Kuang, Milhorn, & Yi, 2016; see Gilroy & Hantula, 2018, for further arguments in favor of model-based AUC).

In this study of 7- to 11-year-olds, as well as using a delay choice task with real rewards of fixed values, we used a more complex delay discounting task. It is well-established that performance in delay of gratification tasks improves developmentally from around 4 years and continues to improve during the school years and across adolescence (e.g., Marchetti et al., 2014; Steinberg et al., 2009). Amongst school-aged children, developmental improvements are evident in both simple delay choice tasks (e.g., Marchetti et al., 2014; Mischel & Metzner, 1962) and in delay discounting tasks (e.g., Steinberg et al., 2009). However, although a number of studies have used the delay discounting task with children before the adolescent period, use of this task with children needs careful consideration, given the relatively complex nature of the task and the fact that the rewards are monetary and hypothetical (Staubitz et al., 2018). Relevant findings come from Burns, Fay, et al. (2019), who examined discounting functions in children aged 7–9 years using a task involving hypothetical monetary rewards. Their procedure involved initial pretraining plus both visual and auditory cues and check questions to ensure children understood the task. After removing children from analysis who produced unsystematic data (Johnson & Bickel, 2008) they were able to derive model-based AUC for children of this age range. Their findings indicate that it is possible to use a child-friendly delay discounting tasks from around 7–8 years.

Of course, model-based AUC scores do depend on the set of candidate functions selected for fitting. Rather than fit an exhaustive list of putative functions, we have selected the three discount functions depicted in Figure 1 on the basis that they are exemplars in the literature and that they make distinct predictions about the discounting of delayed rewards that can be discriminated with a relatively simple delay discounting task that employs a limited number of delays. An exponential discount function implies a constant discount rate over time, and consistent temporal preferences (Halevy, 2015), giving it the status of a normative account of discounting. For a hyperbolic discount function, by contrast, the discount rate is inversely related to the delay to receipt of the reward (Mazur, 1987). Studies with adults have tended to find that, in comparison with exponential discounting, the hyperbolic discount function is a better descriptor of how adults’ valuations of delayed rewards changes over time (Kirby & Herrnstein, 1995). Finally, the Beta-Delta function combines elements of both hyperbolic and exponential discounting (Laibson, 1997). Like the hyperbolic function, it implies increasing patience at longer delays; however, it does so by the inclusion of a second parameter (the beta parameter) to the exponential discount function (the delta parameter) that captures present bias, a prominent feature of people’s temporal preferences.

It should be noted that the central analyses reported in the results section of this study focus simply on group and individual differences in model-based AUC; our primary interest in how this measure relates to other variables. Further details associated with model-fitting are given in the Supporting Information.

Future Time Perception and Delay of Gratification

What might explain differences in children’s ability to delay gratification? One possible factor may be the perception of the length of the delay between now and the point in the future at which the LLR will be available, that is, how far away points in time in the future feel (Bradford, Dolan, & Galizzi, 2019; Han & Takahashi, 2012; Kim & Zauberman, 2009; Kim, Zauberman, & Bettman, 2012; Zauberman, Kim, Malkoc, & Bettman, 2009). Kim and Zauberman influentially argued that intertemporal choices should be considered not as a function of objective future time (actual distance of a time point away from the present) but of subjective future time (how far from the present it feels as though a time point is). Objective and subjective future time may not have a linear relation. Thus, for example, although, objectively, a year from now is 52 times farther away than 1 week from now, this may not feel subjectively to be the
case. In fact, under some circumstances, or for some participants, points in time that are in reality fairly close to the present may feel as though they are almost as far away as points some distance away. This raises the possibility that steep discount rates over short delays might result from a skewing of subjective future time in this way.

Kim and Zauberman (2009; Zauberman et al., 2009) examined whether there is indeed such a non-linear relation between objective and subjective future time. In a number of studies, they elicited a series of judgments about the subjective distance of a variety of time points in the future and showed that a power function (Equation 1) could be used to capture how subjective time ($T$) changes with objective delays ($t$).

$$T = \alpha t^\beta.$$  \hspace{1cm} (1)

Equation 1 has two free parameters, an exponent, beta, that captures the degree of compression in future time perception and a coefficient, alpha. While we refer to the exponent of the power function as capturing the degree of compression in future time perception, others refer to nonlinearity (Han & Takahashi, 2012; Kim & Zauberman, 2009) or to increasing deceleration (Agostino, Zana, Balci, & Claessens, 2019). The alpha parameter is a scaling factor restricted to any positive number that is responsible for the overall height of the time function (see Figure 2a). Kim and Zauberman (2009) claim that the alpha parameter is not merely an arbitrary scaling factor but rather captures individual differences in how far away future time points are perceived to be. The beta exponent is permitted to vary in the range from 0 to 1. When the value of beta approaches 1, the relation between subjective time and objective time approximates to linear. As the beta value diminishes subjective time judgments at longer delays become increasingly compressed relative to objective time (e.g., a week away might feel like $\frac{1}{10}$ of a year away rather than only $\frac{1}{52}$ of a year); see Figure 2b for an illustration of how changes in beta affect the relation between objective and subjective time. Studies with adults have estimated the value of beta to vary within the range .72–.87, indicating a moderate time compression (Agostino et al., 2019; Kim & Zauberman, 2009).

Evidence for this approach to performance on delay discounting tasks comes from studies that show that manipulations that affect future time perception also have parallel effects on patterns of performance on a delay discounting task (Kim & Zauberman 2013; Kim et al., 2012; Zauberman et al., 2009). Moreover, Kim and Zauberman (2009) showed that individual differences in the extent to which adult participants discounted future rewards were correlated with the alpha and beta parameters given in Equation 1, with larger alpha values and smaller beta values associated with steeper discounting. Thus, they argued that

While [previously demonstrated] individual and group differences in temporal discounting have often been attributed to individual’s impulsive reactions to immediate (vs. delayed) outcomes, we suspect that many of these findings can be explained, at least in part, by the differences in time perception. (p. 99)

This raises the possibility that perception of future time is developmentally significant with
regard to delay of gratification: it could be that children’s difficulties in choosing larger later rewards stem at least in part from the way they perceive the intervening time period. That impaired time perception may contribute to impulsive decision making has been explored in children with attention deficit hyperactivity disorder (ADHD; e.g., Rubia, Halari, Christakou, & Taylor, 2009; Toplak, Dockstader, & Tannock, 2006). However, these studies have involved examining duration judgments for periods of time in the order of milliseconds to seconds that children have actually experienced, and moreover the hypothesis that distorted duration perception over these time scales can explain impulsive choice in ADHD has had only limited support (Sonuga-Barke, Bitsakou, & Thompson, 2010; Wittmann & Paulus, 2008). To our knowledge, no previous published work has examined the nature of children’s future time perception (though see Burns, McCormack, et al., 2019). Thus, it is not known whether there are similar functions relating subjective to objective future time in children as there are in adults, or whether the parameters in these functions relate to children’s performance on delay discounting tasks.

Episodic Future Thinking

A further aspect of cognition that has long been considered relevant to the valuation and discounting of distant rewards is future prospection (Böhm-Bawerk, 1889). In more recent times, Boyer (2008) argued that EFT may have evolved to aid future-oriented decision making by making apparent the emotional significance and thus the value of future rewards. Boyer’s thesis has prompted a succession of studies with adults examining whether cueing EFT prior to making intertemporal decisions increases patient responding (Benoit, Gilbert, & Burgess, 2011; Bromberg, Lobatcheva, & Peters, 2017; Bulley et al., 2019; Dassen, Jansen, Nederkoorn, & Houben, 2016; Peters & Büchel, 2010). These studies have consistently revealed a beneficial effect of EFT cueing over and above a variety of control conditions. The majority of these studies involve short-lived “state” changes in discounting behavior, however, the link between EFT cueing and reduced discounting of delayed rewards suggests that “trait” differences in EFT may be an important factor in delaying gratification. This is of particular interest to developmental psychologists, because EFT is known to have a protracted development in early and middle childhood, with adult-like levels of prospection not evident until early adolescence in some studies (Coughlin, Lyons, & Ghetti, 2014; Gott & Lah, 2013). This suggests that developmental improvements in EFT may be at least in part responsible for developmental changes in the ability to delay gratification.

As things stand, there is mixed evidence from correlational studies to support this suggestion. Atance and Jackson (2009) examined EFT using a variety of measures alongside delay of gratification in preschoolers, but found no relation between these abilities when controlling for verbal abilities. More promising findings have come from two studies with adolescents. Bromberg, Wiehler, and Peters (2015) reported a significant association between the amount of episodic details produced in response to future thinking prompts by a sample of 12- to 16-year-olds and a measure of performance on a standard delay discounting task. This relation was significant after controlling for a number of variables including age, IQ, episodic memory, and impulsivity. McCue and colleagues in a study with a much larger sample also found that EFT predicted adolescent discounting (McCue, McCormack, McElnay, Alto, & Feeney, 2019). However, to our knowledge, no studies have examined the relation between EFT skills and intertemporal choice in middle childhood.

One possible explanation of the difference between the correlational findings of Atance and Jackson (2009) with preschoolers and those with adolescents is that when EFT is first beginning to emerge in the preschool years, it is not yet sufficiently established to facilitate delay of gratification. An alternative possible explanation may lie in the differences between the types of tasks used to measure the delay of gratification in preschoolers compared to adolescents, with preschoolers making choices about real rewards that were visibly available (stickers) and adolescents making judgments about hypothetical monetary rewards. It is possible that EFT skills are particularly important when dealing with hypothetical rewards, perhaps due to the involvement of imagination. Indeed, in adult EFT priming studies, the majority of intertemporal choice trials are unrewarded, with perhaps a single trial chosen at random after the task has been completed to determine the reward for the participant (though see Daniel, Said, Stanton, & Epstein, 2015; Daniel, Stanton, & Epstein, 2013; Dassen et al., 2016). As things stand, we are not able to decide whether either of these two possible explanations of the differences in findings from the preschooler study and the adolescent studies is correct.
Current Study

In this study, we examined delay choice, delay discounting, future time perception and EFT, alongside verbal and nonverbal intelligence, in a large sample of 7-to-11-year-olds. The delay choice task involved similar delay periods to the delay discounting task, but with real rewards and fewer trials. The child-friendly delay discounting task involved hypothetical monetary rewards. Administering this set of tasks allowed us to address a number of issues. First, it allowed us to examine whether performance on a simple delay choice task and delay discounting are related in middle childhood. Although it is often assumed that they measure similar abilities, Burns, Fay, et al. (2019) found that performance on two versions of these tasks were not correlated when controlling for verbal and nonverbal ability. Second, we explored, for the first time, whether children showed a similar power function (Equation 1) relating objective future time to subjective future time as adults by using a novel future time perception task that we believed to be suitable for use with children. Third, we were able to explore whether children’s performance on the future time perception task was related to their performance on either of the intertemporal choice tasks, which would indicate a role for future time perception in explaining delay of gratification. Fourth, we also examined the relations between the two measures of intertemporal choice and children’s EFT abilities, as given by their performance in a future thinking interview. If EFT’s function is to aid future-oriented decision making then we might expect to see children with better EFT skills behave more prudently on intertemporal choice tasks. The inclusion of the two intertemporal choice tasks allowed us to test for an association between EFT and performance on tasks involving real and hypothetical rewards, given the possibility that EFT may primarily be related to performance on tasks involving hypothetical rewards.

As we were collecting a very large sample of EFT data from this age group, which has been relatively understudied with regard to EFT (though see Coughlin et al., 2014), we also followed Coughlin et al. in taking measures of participants’ ratings of the clarity of the imagined future events, the speed with which these came to mind, and their valence. This allowed us to look at how these varied as a function of the distance in the future of the to-be-imagined event. Finally, we included IQ measures as a covariate in our individual differences analyses as previous studies have indicated that higher intelligence is associated with increased patience on delay of gratification tasks and delay discounting tasks (Bromberg et al., 2015; Shamosh & Gray, 2008).

Method

Participants

We recruited participants from three age groups according to the school year group in which they were enrolled: 7- to 8-year-olds (fourth year of formal education in the UK), 9- to 10-year-olds (sixth year of formal education in the UK) and 10- to 11-year-olds (seventh year of formal education in the UK). Based on similar studies in the literature (Atance & Jackson, 2009; Hanson, Atance, & Paluck, 2014) we aimed to recruit 40 participants per age group. Data collection was stopped whenever a minimum of 40 participants in age group had completed the study and when all the children had been tested from a given school class whose parents or guardians had provided written consent. The final sample consisted of 132 children (64 male): forty-nine 7- to 8-year-olds ($M_{age} = 7$ years 11 months; $SD = 3.8$ months; range = 89–101 months); forty-one 9- to 10-year-olds ($M_{age} = 9$ years 10 months; $SD = 3.8$ months; range = 112–126 months); and forty-two 10- to 11-year-olds ($M_{age} = 11$ years 1 months; $SD = 3.7$ months; range = 126–139 months). One child from the 9- to 10-year-old group did not provide data on the delay discounting task or future time perception task due to absenteeism on the second day of data collection. Participants were recruited from three schools in Belfast, Northern Ireland. The schools drew largely from an urban population of low to middle socioeconomic status and, due to the demographics of the population in Northern Ireland which is approximately 98% white, the vast majority of sample was white. Participants were all native speakers of English. Informed written consent for all tasks was sought from parents or guardians prior to testing. Parents or guardians of nine participants did not consent to the recording of EFT interviews and so this task was not conducted with these children. The testing took place between November 2017 and February 2018.

Measures

Delay Choice

This task comprised 12 trials on which participants were offered a choice between a low value reward available immediately and a higher value
reward available after a specified delay. Three delay periods were used, 1 day, 1 week, and 1 month, with four trials at each delay. On each trial, one of four reward pairs (once at each delay) was offered: trading cards (1 vs. 2), erasers (1 vs. 2), pens and pencils (1 vs. 2), and sweets (1 vs. 2).

Participants sat at a table across from the experimenter, with two trays, one labeled “now” and one labeled “later,” placed in front of them. The experimenter showed children three types of trading card, each from a different theme (Football players, Star Wars characters, and Despicable Me cards for the males; Disney Princess, Shopkins, and Trolls cards for the females) and two types of novelty erasers (Lego shaped and emoji). Each child’s preferred trading card and eraser was used in the subsequent delay choice trials. On trials on which the trading card was the reward the cards were placed inside a small opaque envelope to hide the characters on the cards, ensuring that the specific cards used did not affect children’s choices.

On each trial, children heard the following: “I have some [reward type] here. I’m going to give you a choice. You can have this one [reward type] right now, or you can have 2 [reward type] in a [delay length] time. If you take this 1 [reward type] here you can get it right now, but you won’t get the other 2 [reward type]. If you want to wait for the two [reward type] you won’t get any right now, but you will get these two [reward type] when I come back [delay length]”. Thus, the delay choice task employed an ‘explicit zero’ framing whereby the participant is reminded of the null outcome entailed by each choice. We note that zero framing has been shown to reduce discounting in adult participants, but our aim in using such framing was to ensure that children fully understood and appreciated the choices available to them rather than to reduce discounting. Children were then asked two check questions: “If you choose this 1 [reward type], when will you get it?” and “If you choose these two [reward type], when will you get them?” If they answered either one of these questions incorrectly the choice and the delays were restated, and the check questions repeated. They were then asked the test question: “What would you like to do? Point to the tray which shows me which you would like.” Participants were scored a 0 on a trial if they chose the immediate reward, and a 1 if they chose the delayed reward; overall scores ranged from 0 to 12. When children chose to delay gratification, they were either given the reward after the delay had elapsed, or, for practical reasons, once all data collection at their school had been completed, whichever came sooner.

Episodic Future Thinking Interview

This task was based on the episodic thinking interview used by Coughlin et al. (2014). Children were asked to describe specific personal events located at one of three temporal periods (tomorrow, next week, and a few months), though unlike the Coughlin et al. study we did not provide children with a cue word on each trial to help generate an event. We used cueing by time period to ensure that the EFT measure involved describing events from time periods that were also included in the delay choice and or the delay discounting task. All children completed the task in the same order (with cues in the order tomorrow, next week, a few months), and they were asked to generate two events for each temporal period. Prior to generating an event for a given period, the experimenter described a particular event that she or he would be carrying out at that period in order to give an example of what children were required to do; for tomorrow the event was about going shopping for a new pair of shoes; for next week, it was about playing a game of football in the park, whilst for a few months, the event involved going to the seaside. Children were prompted to give a specific rather than continuous event (“Can you tell me about something that is going to happen to you just one time tomorrow/next week/in a few months from now?”). Children were prompted for further details using the same prompts outlined in Coughlin et al.: (a) Can you tell me more about what will happen? (b) Can you tell me more about where this will happen? (c) Can you tell me more about who will be there? Each prompt was used once for any given temporal period.

We also asked children to make a series of judgments following their event descriptions. Two of these were similar to those used by Coughlin et al. (2014)—speed with which the event came to mind and clarity of the event—and the third was an emotion judgment. Three specially constructed visual scales were used to elicit each judgment (see Supporting Information for details). The speed and clarity visual scales were similar to those used in Coughlin et al. (2014). Speed judgments were made on a three-point scale ranging from “slowly” (represented by a tortoise and scored as 1) to “quickly” (represented by a hare and scored as 3). Participants were asked, “when you thought about [reported event] how quickly did it come into your head?” A six-point clarity scale was used in which
each point was represented by the figure of a child with a thought bubble containing an image of a farm. The thought bubbles contained an increasingly more detailed image as one went from “not at all clear” (score of 1) to “very, very clear” (score of 6) on the scale. Participants were asked, “when you thought about [reported event] how clear did it look to you in your head?” Finally, participants were asked to indicate how each event made them feel, using a 7-point emotional faces scale ranging from “very sad” (scored as 1) to “very happy” (scored as 7). Participants were asked, “when you thought about [reported event] how did it make you feel?” Participants were given training and instruction on how to use each scale during the experimenter demonstration. In the interview, each scale was presented on a touchscreen laptop, though the questions were presented orally.

**Future Time Perception**

In this task, which was broadly based on that used by Kim and Zauberman (2009), participants were asked to consider how far away eight future time points felt to them (1 day, 3 days, 1 week, 2 weeks, 3 weeks, 1 month, 2 months and 3 months). Kim and Zauberman used a procedure in which participants pulled a line along a computer screen; we made this task more concrete by getting children to pull a length of cord from a specially constructed dispenser (a wooden box of dimensions 160 × 160 × 150 mm). The dispenser housed a large reel of cord. On one side of the box there was an aperture of diameter equivalent to the thickness of the cord, from which the cord could be pulled. A lever on the side of the box allowed dispensed cord to be wound back into the box.

Children were given the following instructions: “In this game I want you to think about how far away some times are in the future. I’m going to ask you to think about how far away tomorrow feels to you and how far away 3 months from now feels to you and how far away a number of times between tomorrow and 3 months from now feels to you. You are going to use this string to show me how far away times in the future feel to you. I’ll show you how it works”. The experimenter gave two demonstrations. For the first demonstration she or he informed participants that she or he was thinking about a “long time from now” and pulled out 360 mm of cord approximately (2 revolutions of the winding handle). Participants then completed eight trials in one of four counterbalanced orders (two of which began with the shortest judgment [1 day] and 2 of which began with the longest judgment [3 months]). The test question was as follows: “I want you to think about [tomorrow/3 days from now/etc.]. How far away does that feel to you? Pull the string to show me how far away [tomorrow/3 days from now/etc.] feels to you.” Once children had pulled out the string, the experimenter then cut off the length that had been pulled out for that trial and stored it before proceeding to the next trial. The length of each string was subsequently measured and recorded to the nearest 5 mm.

**Delay Discounting**

This was a computerized delay-discounting task using hypothetical rewards. The delayed reward was £10 (UK pounds) and offered at one of four delays: 1 day, 1 week, 2 weeks, and 3 months. The immediate reward ranged from 0 to £10 and varied in 50p increments. We employed an iterative optimizing procedure to hone in on indifference points at each delay (Richards, Zhang, Mitchell, & de Wit, 1999). The value of the immediate reward on each trial was determined by an adjusting procedure based on each participant’s choice on the previous trial at that same delay. Full details of the adjusting procedure are given in the Supporting Information. Across the whole sample optimization was reached after a mean number of 47 trials (range = 27–81).

The task was programmed in E-prime Software (Psychology Software Tools, Pittsburgh, PA) and administered on a 15 in. Dell laptop with touchscreen. Participants were told that they would be given a series of choices between two sums of money, one available immediately and one available after a delay. They were instructed to select the option that they preferred. Figure 3 shows the structure of each trial. Rewards and associated delays were presented serially each with an accompanying audio description of the value of the reward and the delay. Participants responded by tapping the picture of their preferred option.

**WISC–IV Subtests**

The vocabulary and block design subtests of the Wechsler Intelligence Scale for Children (4th ed.; WISC–IV; Wechsler, 2003) were administered. The raw scores from these tests were used in the analysis.
Procedure

Due to the number of measures, children completed the study in two sessions, with the delay choice task and the EFT interview in the first session. The next session on the following day involved the delay discounting task, future time perception task and two subtests of the WISC–IV. We employed a fixed task order as we were primarily interested in individual differences as the source of variation across the tasks. A fixed task order minimizes order effects as a source of noise (Goodhew & Edwards, 2019).

Data Processing

Delay Discounting

The data were first subjected to a systematicity criterion (Johnson & Bickel, 2008). Unsystematic discounters were those that produced an indifference point for a given delay period that was greater in value than the preceding delay period by more than £2 (i.e., an increase > 20% of the LLR). These participants were removed from all analyses concerning delay discounting. For the remaining participants, their delay discounting data were analyzed in a two-step process. First, a number of discount functions were fitted to each participant’s data. These discount functions were selected based on their normative (exponential discount function) or putative descriptive status (hyperbolic and quasi-hyperbolic discount functions). A fourth “noise” model, a y-intercept model based on the mean indifference point that captures insensitivity to variations in delay was also fitted. The best-fitting model for each participant was determined on the basis of Bayesian information criterion scores (BIC; Schwarz, 1978). In step two we derived model-based AUC (model AUC), by applying integral.
calculus to the fitted function (Gilroy & Hantula, 2018). Model AUC scores are scaled to the interval 0–1 by dividing them through by the total area of the graph.

**Delay Choice**

Similar to the approach recommended by Johnson and Bickel (2008) for delay discounting data, we applied a novel criteria to delay choice scores to identify nonsystematic discounting on the task by comparing participants’ delay scores on consecutive delay lengths. If at the 1-week or 1-month delay length children’s delayed choice total score was greater by 2 or more than their delay choice total score for the 1-day and 1-week delays, respectively, then that participant’s data were considered unsystematic and they were removed from analysis concerning delay choice. So for example, if a participant’s delay choice total for the 1-day delay was 0 they would be considered unsystematic if their delay choice total for the 1-week trials was 2 or more.

**Future Time Perception**

For the future time perception task, we first identified participants who failed to produce estimates that monotonically increased with an actual distance of the time point in the future. Estimates were compared for successive time points: We classified as unsystematic responders those who produced two or more time estimates that decreased by > 20% over the estimate at the previous time point. Data from unsystematic responders were not included in the analyses that involved this task. For each participant, we first transformed future time estimates into subjective units by setting the median string length for the shortest 1-day estimate equal one unit of subjective time. All time estimates were transformed into subjective units by dividing them by this number.

**EFT Interviews**

Transcribed EFT interviews were coded using a scheme owing to Levine, Svoboda, Hay, Winocur, and Moscovitch (2002), and previously used by Bromberg et al. (2015) and McCue et al. (2019). The transcriptions were first segmented into unique details, that is, simple subject predicate clauses. Each clause was then categorized as either an internal detail or an external detail. Internal details were those that pertained to the main event being discussed and belonged to one of five mutually exclusive subcategories: event detail, time, place, perceptual, and emotion or thought. External details were details pertaining to events other than the identified main event. Semantic details that did not require recollection of a specific time and place and repetitions of internal details were also coded as external. To check for reliability a second independent coder scored 30% of the transcripts. Intraclass correlations (ICC; two-way mixed effects) were high for internal details (ICC = .79, p < .001). The average internal detail score across each pair of episodes at each delay was calculated.

**Data Analysis Plan**

Data analysis focused on examining how the factors of age and delay impacted on performance on delay of gratification, delay discounting and EFT, followed by correlational and regression analyses to examine the associations between the measures, while controlling for age and IQ. Analyses were for the most part planned prior to the commencement of data collections and tended toward confirmatory in nature in so far as there were clear directional hypotheses for each analysis. Mixed analysis of variance (ANOVA) was planned to examine the effect of age group (7- to 8-year-olds, 9- to 10-year-olds and 10-to 11-year-olds) and delay (1 day, 1 week, 1 month) on delay choice scores and EFT scores, whereas a one-way ANOVA assessing the effect of age group on model-based AUC scores was planned. These analyses were carried out because findings from previous studies indicated that both age and delay might be expected to influence performance on all three tasks, with performance improving as a function of age and were therefore confirmatory in nature. The selection of models and model fitting methodology (ordinary least squares), for both the delay discounting and time estimation task was planned in advance of data collection. The application of the systematicity criteria for the delay discounting task was planned in advance of data collection and based on pre-existing systematicity criteria used in other studies; however, the systematicity criteria for the time perception task and the delay choice task were developed after the data had been collected. Finally, analyses of individual differences were planned using correlations and regression analyses. The model fitting for the delay discounting task was done using the Discount Model Selector software (Gilroy, Franck, & Hantula, 2017), whereas the model fitting in for the time estimation task was
done in Excel using the Solver add-in program. All remaining analyses were completed in R (R Core Team, 2015).

Results

A series of t-tests revealed no effect of gender on delay of gratification totals, delay discounting switch points, future time judgments, mean internal detail scores, and IQ measures (all p values > .13); gender was not considered further in the analyses.

Delay Choice (Real Rewards)

Applying the systematicity criteria removed 12 participants from the sample (seven 7- to 8-year-olds, two 9- to 10-year-olds, and three 10- to 11-year-olds). Mean numbers of delayed choices for each delay and age group are shown in Figure 4. A two-way mixed ANOVA on delay of gratification scores with delay (1 day, 1 week, and 1 month) as a within subjects factor and age group (7- to 8-year-olds, 9- to 10-year-olds, 10- to 11-year-olds) as a between subjects factor revealed a significant main effect of delay, \( F(2, 234) = 79.81, p < .01, \eta^2_G = .17 \), a main effect of age group, \( F(2, 117) = 6.10, p < .01 \), \( \eta^2_G = .07 \), and a significant interaction between delay and age group, \( F(4, 234) = 5.07, p < .01, \eta^2_G = .03 \). Three follow-up one-way analyses of variance, one for each delay period, were conducted with age group as a between subjects factor. There was a significant main effect of age group at 1-day delay, \( F(2, 117) = 11.30, p < .01, \eta^2_G = .16 \), and a significant linear trend, \( F(1, 117) = 21.081, p < .01 \), indicating that older age groups delayed more often. The main effect of age group, \( F(2, 117) = 7.36, p < .01, \eta^2_G = .11 \), and linear trend, \( F(1, 117) = 14.53, p < .01 \), was also found for the 1-week delay. However, there was no effect of age group on delay at 1 month, \( F(2, 117) = 0.19, p = .83 \).

Delay Discounting

Applying the systematicity criterion removed 31 participants from analyses involving delay discounting. Data from two further children who selected the immediate reward on every single trial, including trials on which the immediate reward was zero, were also removed. Participants removed from the analyses were significantly younger than those who were retained (106 months vs. 117 months, \( t(56.9) = 3.47, p < .001 \)). The remaining children represented 75% (\( N = 98 \)) of the whole

![Figure 4. The mean number of delayed choices by age group and delay period. Error bars represent standard error.](image-url)
sample. As described earlier, we derived model-based AUC; details are provided in the Supporting Information of the distributions of children’s responses across the different functions, however, we note here that the exponential discount function proved the best fitting model for almost half the participants ($n = 42$). A one-way ANOVA of model-based AUC with age group as a between subjects factor revealed no significant effect of age group, $F(2, 95) = 0.30, p = .74$.

**Episodic Future Thinking**

Sixteen children (13%) failed to produce an event for at least one of the six trials during the EFT interview. Where they provided at least one episode for each delay ($n = 12$), their score for that one episode was taken as the mean delay score. The remaining four children were removed from analyses involving EFT. Data from one further participant are missing due to recording equipment failure. A repeated measures ANOVA examined the effect of delay (1 day, 1 week, and a few months), age group (7- to 8-year-olds, 9- to 10-year-olds, and 10- to 11-year-olds) and the interaction of delay and age group on the internal detail scores. There was a main effect of delay only, $F(2, 230) = 10.54, p < .01, \eta^2_p = .03$. Table 1 displays the mean internal detail scores for each delay collapsed across age group. As can be seen from the table, internal detail scores actually rose as delay length increased. Paired samples t-tests revealed that tomorrow events produced significantly fewer internal details than next week events, $MD = 0.67$, 95% CI [0.18, 1.17], $t(118) = 2.71, p < .01$, and than events a few months away, $MD = 1.19$, 95% CI [0.65, 1.73], $t(118) = 4.36, p < .001$. Internal details scores for next week were also significantly lower than for events a few months away, $MD = 0.51$, 95% CI [0.01, 1.02], $t(118) = 2.01, p = .047$. Cronbach alpha scores indicated good internal reliability for internal details ($\alpha = .79$) so an averaged measure over the three delays was considered in the individual differences analysis.

Delay length had a significant effect on ratings of how quickly events came to mind, $F(2, 234) = 4.09, p = .018, \eta^2_p = .02$. Events a few months away were brought to mind significantly quicker than next day events, $MD = 0.18$, 95% CI [0.06, 0.31], $t(117) = 2.84, p < .01$. There was a main effect of delay on emotion ratings, $F(2, 234) = 3.65, p = .028, \eta^2_p = .02$. Next week events were rated more positively by participants than tomorrow events, $MD = 0.30$, 95% CI [0.09, 0.51], $t(117) = 2.80, p < .01$. None of the other comparisons were significant ($p > .1$). However, speed, clarity and emotion ratings showed poor internal reliability (all $\alpha < .50$). These variables were not considered further in the individual differences analysis.

**Future Time Perception Task**

Applying the systematicity criteria removed 16% ($N = 21$) of the sample from further analysis. There was no difference in the mean age of those removed from the analysis (111 months) and those retained (115 months), $t(25.9) = 0.82, p = .42$. Those removed from the analyses did have lower Block Design scores (21.1 vs. 27.1, $t(28.9) = 2.20, p < .05$). We modeled the relation between calendar time and subjective time estimates by fitting both linear $f(t) = mt + b$ and power functions $f(t) = at^b$. Models were fitted in R using the lme4 package (Bates, Maechler, Bolker, & Waler, 2014; R Core Team, 2015) to account for the clustered nature of the data (model parameters were varied by participant). BIC scores (lower scores signal better model fit) indicated that the power function provided a better fit for the data than the linear function (BIC power model = 3,160.8, BIC linear model = 3,291.5). We computed a Bayes factor (assuming flat priors) to assess the relative strength of the evidence in favor of the power model over the linear model. The Bayes factor was $1.87^{28}$: for reference Bayes factors > 150 are considered very strong evidence. Figure 5 below displays the power model. The model parameter values were 1.20 and 0.30 for alpha and beta, respectively, indicating a high degree of future time compression. To illustrate this, the model estimates that, for children, the shortest period of 1 day in the future feels on average to be about $\frac{1}{3}$ of the longest period of 3 months (whereas in reality is about $\frac{1}{6}$). For individual differences analyses, we fitted individual power functions to each child’s time estimates deriving $\alpha$ and $\beta$ parameter values.

---

Table 1: Mean Episodicity Scores and Associated Speed of Recall, Clarity, and Emotion Ratings for Recalled Events in the Episodic Future Thinking Interview

<table>
<thead>
<tr>
<th>Event</th>
<th>Internal details</th>
<th>Speed</th>
<th>Clarity</th>
<th>Emotion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tomorrow</td>
<td>8.06 (2.74)</td>
<td>2.06 (0.53)</td>
<td>4.53 (0.98)</td>
<td>5.81 (1.04)</td>
</tr>
<tr>
<td>Next week</td>
<td>8.73 (3.17)</td>
<td>2.20 (0.52)</td>
<td>4.68 (0.78)</td>
<td>6.11 (0.80)</td>
</tr>
<tr>
<td>Few months</td>
<td>9.25 (3.12)</td>
<td>2.25 (0.59)</td>
<td>4.58 (1.23)</td>
<td>6.00 (1.10)</td>
</tr>
</tbody>
</table>
Relations Between Measures

We first examined the correlations between demographic measures (age and IQ scores) and measures of delay choice, delay discounting (model-based AUC), EFT and future time perception (alpha and beta parameters). Table 2 below presents zero-order correlations above the diagonal and partial correlations controlling for age and the IQ measures below the diagonal. There was a significant correlation between children’s scores on the delay choice task and model-based AUC, which remained significant after controlling for age and IQ scores. Delay choice and model-based AUC scores were, however, differentially related to EFT and future time perception measures. AUC was positively correlated with both EFT and beta parameter values, indicating that steeper discounting is associated with lower EFT and greater time compression. However, only the correlation with beta remained significant after controlling for age and IQ effects. Delay choice tasks on the other hand were unrelated to EFT and future time perception measures.

To explore the relation between delay discounting scores and both future time perception and EFT we conducted a hierarchical multiple linear regression with model-based AUC as the dependent variable. An initial model included age, WISC vocabulary scores and WISC block design scores as predictor variables. In comparison to the null model this model was a significant predictor of model-based AUC, F(3, 72) = 2.82, p = .045. The model had an $R^2$ value of .105 (adjusted $R^2 = .068$). In a second step beta and EFT scores were added as predictors. This model had an $R^2$ value of .177 (adjusted $R^2 = .117$). An ANOVA comparing this model to the model with just age, vocabulary, and block design as predictors indicated a marginally improved model fit, F(2, 70) = 3.03, p = .055, $\Delta R^2 = .072$. Beta and EFT were therefore retained in the final model, presented in Table 3 below. Beta scores were the only significant predictor in the model.

Discussion

This study is the first to examine the relation between EFT, future time perception, and performance on intertemporal choice tasks in middle childhood. We used a delay choice task with real rewards and a delay discounting task with hypothetical monetary rewards, allowing us also to examine the relation between performance on these two quite different types of delay of gratification tasks. To the best of our knowledge, this study is
the largest and most systematic to date to simultaneously examine children’s performance on both these types of tasks. We found that there was a significant relation between these two tasks, and that this remained the case even when controlling for age and IQ. Our novel future time perception task also allowed us to look for the first time at the relation between objective and subjective future time perception in children. This task yielded meaningful data, and, as with adults, the relation between objective and subjective future time was best captured by a power function that indicated a nonlinear relation. We also found that, of all our measures, the degree of future time compression was the best predictor of children’s performance on the delay discounting task, although it did not relate to performance on the delay choice task. Children’s ability to generate episodic descriptions of events in the distant future was also significantly related to performance on the delay discounting task, but this relation did not survive controlling for age and IQ. We now consider each of these findings in more detail in the light of existing theoretical and methodological debates about delay of gratification and its development, although we note that the generalizability of our findings is constrained by demographics of our sample (primarily white urban children), with recent research suggesting that the relation between delay discounting performance and other measures may differ in other populations (Duran & Grissmer, 2020).

**Delay Choice and Delay Discounting**

The older children were more likely to delay gratification on the delay choice task than the younger children, at least for the shorter delays, but we did not find age-related changes on the delay discounting task. Nevertheless, we did find that performance on these two intertemporal choice tasks was significantly related, and, unsurprisingly in the light of previous research suggesting a link between delay of gratification and intelligence (Shamosh & Gray, 2008), performance on both tasks was also related to the IQ measures. We note that Burns, Fay, et al. (2019) found that performance on a delay choice task and a delay discounting task was unrelated in a sample of children of a similar age range. However, their delay choice involved fewer trials and only one short delay (a day), whereas in this study our delay choice involved some larger delays that were of a similar magnitude to those used in the delay discounting task as well as a greater number of trials.

Delay discounting studies with adults have examined whether using real versus hypothetical rewards makes a difference to patterns of performance (e.g., Johnson & Bickel, 2002; Lokey, Jones, & Rachlin, 2011). While these studies show mixed findings with regard to whether the nature of the

---

Table 2
Zero-Order Correlations Are Displayed Above the Diagonal and Partial Correlations Controlling for Age and IQ Measures Are Displayed Below the Diagonal. Exclusions Are Performed Pairwise, Therefore N Varies Across Cells (83–132)

<table>
<thead>
<tr>
<th></th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
<th>5.</th>
<th>6.</th>
<th>7.</th>
<th>8.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Age</td>
<td>.48***</td>
<td>.47***</td>
<td>.30***</td>
<td>.09</td>
<td>.11</td>
<td>.10</td>
<td>.04</td>
</tr>
<tr>
<td>2. Vocabulary</td>
<td>—</td>
<td>.51***</td>
<td>.35***</td>
<td>.24*</td>
<td>.20</td>
<td>.10</td>
<td>.04</td>
</tr>
<tr>
<td>3. Block design</td>
<td>—</td>
<td>—</td>
<td>.42***</td>
<td>.20</td>
<td>.10</td>
<td>.04</td>
<td>.01</td>
</tr>
<tr>
<td>4. Delay choice</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.45**</td>
<td>.14</td>
<td>.03</td>
<td>—</td>
</tr>
<tr>
<td>5. Model-based area under the curve</td>
<td>.41***</td>
<td>—</td>
<td>.21*</td>
<td>.05</td>
<td>.26***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Episodic future thinking</td>
<td>.09</td>
<td>.18</td>
<td>—</td>
<td>.07</td>
<td>.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. α</td>
<td>.02</td>
<td>.05</td>
<td>—</td>
<td>.09</td>
<td>—</td>
<td>—</td>
<td>.37***</td>
</tr>
<tr>
<td>8. β</td>
<td>—</td>
<td>.24*</td>
<td>.12</td>
<td>—</td>
<td>.37**</td>
<td>—</td>
<td></td>
</tr>
</tbody>
</table>

*p < .05,  ** p < .01,  *** p < .001.

Table 3
Multiple Regression Model Predicting Model-Based Area Under the Curve Score on the Delay Discounting Task (N = 76)

<table>
<thead>
<tr>
<th>Variable</th>
<th>B (SE)</th>
<th>t</th>
<th>p</th>
<th>95% CI for B</th>
<th>95% CI for B</th>
</tr>
</thead>
<tbody>
<tr>
<td>(intercept)</td>
<td>-0.03 (.26)</td>
<td>-0.11</td>
<td>.91</td>
<td>-0.54, .49</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>.00 (.00)</td>
<td>-0.29</td>
<td>.78</td>
<td>-0.01, .00</td>
<td></td>
</tr>
<tr>
<td>Vocabulary</td>
<td>.01 (.00)</td>
<td>1.75</td>
<td>.08</td>
<td>.00, .02</td>
<td></td>
</tr>
<tr>
<td>Blocks</td>
<td>.00 (.00)</td>
<td>.47</td>
<td>.64</td>
<td>.00, .01</td>
<td></td>
</tr>
<tr>
<td>Episodic future thinking average</td>
<td>.01 (.01)</td>
<td>1.00</td>
<td>.32</td>
<td>-0.01, .04</td>
<td></td>
</tr>
<tr>
<td>β</td>
<td>.37 (.18)</td>
<td>2.10</td>
<td>.04</td>
<td>.02, .72</td>
<td></td>
</tr>
</tbody>
</table>
reward affects the nature of discount curves, they consistently show robust correlations between performance on tasks involving real and hypothetical rewards (e.g., Hinest & Anderson, 2010; Matusiewicz, Carter, Landes, & Yi, 2013). Thus, our results align in this respect with findings from the adult literature. In the adult work, studies comparing real and hypothetical rewards have been conducted primarily because researchers have been concerned with whether delay discounting tasks involving hypothetical rewards are a valid means to measure impulsive choice, given that such tasks are easier to administer, both practically and ethically. Although we did find a relation between children’s performance on the delay choice task and the delay discounting task, we would be reluctant to conclude that the delay discounting task can be used as a substitute for the delay choice task in a developmental context. For a start, as we have pointed out, the tasks show differential effects of age in our study. Moreover, as in Burns, Fay, et al. (2019) study, around 25% of children did not produce systematic data in the delay discounting task, leading to their exclusion from analysis. We do note that those removed from the delay discounting task analyses were significantly younger than those included; however, even if they are also removed from the delay choice analyses the correlation between delay choice and age remains significant ($r = .28, p = .005$). Thus, although our findings suggest these tasks draw on some of the same processes, they are not interchangeable.

**Future Time Perception**

Children in our study seemed to quickly grasp what they were required to do in the future time perception task, which essentially involved mapping magnitude between two dimensions (spatial length and future time). That most children produced judgments that were monotonically ordered with respect to objective time indicates that the task has good internal reliability. Previous studies suggest that even infants map magnitudes between spatial extent and temporal duration (de Hevia, Izard, Coubart, Spelke, & Streri, 2014; Lourenco & Longo, 2010; Srinivasan & Carey, 2010). However, the current task involved not simply mapping experienced duration (in the order of seconds) on to spatial length, but children reporting how far away in the future time points felt subjectively. To the best of our knowledge, the only analogous task that has been used in a developmental context is that of Burns, McCormack, et al. (2019), who asked children as young as 4–5 years to judge how far away Halloween felt 2 weeks before or 2 weeks after the event itself. Children made a single judgment by pointing to a location on a line; replicating findings with adults (Caruso, Van Boven, & Chin, 2013), Halloween was judged to feel closer when it was 2 weeks in the future than when it was 2 weeks in the past. This study extends this methodology by demonstrating that it is possible to get children to make a range of judgments about subjective distances in time at least from the age of around 7 years. The findings complement those of Tillman, Marghetis, Barner, and Srinivasan (2017) who demonstrated that from about 7–8 years children were able to make judgments about the objective remoteness of points in the future such as next week and next year.

As in studies with adults (Kim & Zauberman, 2009; Zauberman et al., 2009; though see Agostino et al., 2019), the relation between objective and subjective time was nonlinear and best captured by a power function (Equation 1). Notably, the beta value of the power function that we obtained (.30), which captures the degree of compression in the relation between subjective and objective time, was substantially lower than that previously obtained with adults (.72; Kim & Zauberman, 2009; .75–.87; Agostino et al., 2019). There are a number of methodological differences between the future time perception task in this study and those that have been administered to adults that may account for this difference. The range of durations used in this study (1 day to 3 months) was smaller than that used by Kim and Zauberman (3 months to 3 years). Moreover, the ratio of the shortest to longest delay was considerably larger in this study (1:91.25) than in both the Kim and Zauberman and Agostino et al. studies (1:12). When coupled with a bounded scale for subjective judgments, such as the line bisection task of Zauberman et al. (2009) and Agostino et al. (2019), a large ratio of shortest to longest duration may artificially lower beta values, due to the restrictions on participant’s responses to long durations imposed by the scale. However, in this study we used an unbounded scale, mitigating this problem. It may be, therefore, that there are developmental changes in future time perception, with children showing substantially more compression of future time, although we note that within the age range sampled there was no evidence of a correlation between age and beta values. Without directly comparing the performance of children and adults on the same task it is not possible to know. What we did show, though, which replicated Kim
and Zauberman’s findings with adults, was that degree of compression was correlated with degree of discounting on a delay discounting task.

We note, though, that unlike Kim and Zauberman, we did not find that individual differences in the alpha parameter were related to performance on the discounting task. Kim and Zauberman argued that the alpha parameter is not just an arbitrary scaling parameter but reflects the “overall” amount of time compression. The lack of predictive value of alpha in our study indicates that we need to consider whether this parameter is simply a reflection of the overall amount of string different children tend to pull out in order to make their judgments (and thus an arbitrary scaling parameter), rather than a reflection of the general extent to which they felt that points in the future are far away. Kim and Zauberman’s task differed, in that participants made their judgments by lengthening a line on a computer screen. While it was possible for participants to extend the length of the line beyond a single screen width, the screen width itself may have functioned as a mode of calibration for participants’ judgments that they all shared. Thus, it may have been that under such circumstances differences in absolute line length were indeed reflective of individual differences in the general extent to which the future felt far away. In contrast, there were no constraints on how much string children could pull out, and the extent to which children pulled out a lot versus a little string may have been reflective of other types of individual differences.

Setting aside this interpretive issue regarding the alpha parameter, the finding that the beta parameter, that is, degree of compression, was significantly related to children’s performance on the delay discounting task is a novel one. We interpret this finding as consistent with Kim and Zauberman’s (2009) suggestion that the pattern of choices in the delay discounting task reflects, at least in part, participants’ perception of the length of the delays they are asked to make decisions about. When the relation between objective and subjective time is one of nonlinear compression, short delays are judged as feeling to be a relatively similar distance away to long delays, which may help explain the steepness of children’s delay discounting. Kim and Zauberman (2009, 2019) argue that future time perception functions as an explanation of individual or group differences in delay discounting that is distinct from one in terms of impulsivity per se, and moreover that interventions that have an impact on future time perception can in turn affect intertemporal choice (Kim et al., 2012; Kim & Zauberman, 2013). The current findings that demonstrate a relation between future time perception in children and delay discounting add to what is as yet still a small body of empirical research that supports Kim and Zauberman’s approach (see also Bradford et al., 2019). However, we note that there was no relation between performance on the future time perception choice and children’s performance on the delay choice task, despite the fact that performance on the delay choice and the delay discounting task were themselves related. It may be that the delay choice task lacks sufficient sensitivity to pick up on a link between future time perception and intertemporal choice. In comparison with delay discounting tasks the delay choice task has many fewer trials and examines intertemporal choices across a shorter time frame (1 day to 1 month in the present task, compared with 1 day to 3 months in the delay discounting task).

As things stand, our findings only provide initial evidence for a link between future time perception and delay of gratification on a task involving entirely hypothetical rather than real rewards. Nevertheless, the findings provide motivation for more fully exploring the role of future time perception in explaining developmental changes in delay of gratification. Much is yet still unknown about developmental changes in this aspect of temporal cognition and how any such changes relate to other aspects of cognitive development. For example, children are still learning to fully master and flexibly make use of the clock and calendar system in the age range that we tested (Friedman, 1989; McCormack, 2015), and it is not known how such mastery may in turn have an impact on children’s performance on delay discounting tasks by affecting their perception of future delays. A further intuitively plausible hypothesis that suggests itself in the context of this study is that future time perception might itself be connected to children’s ability to richly imagine time points in their personal futures—that is, engage in EFT. However, we found no evidence of a link between EFT and performance on the future time perception task, and indeed only weak evidence of a link between EFT and intertemporal choice, a finding we now turn to discussing.

**EFT and Intertemporal Choice**

Evidence in our study for a link between EFT and intertemporal choice was weak. There was no correlation between EFT and performance on the delay choice task, and the significant correlation between EFT and performance on the delay discounting task did not survive controlling for age
and IQ. These findings contrast with those with adolescents previously reported by Bromberg et al. (2017) and McCue et al. (2019), which both found that EFT was a predictor of performance on a delay discounting task, although they are consistent with Atance and Jackson’s (2009) findings with preschoolers. These previous findings were taken as evidence for accounts that suggest that imagining future events reduces impulsive behavior (Benoit et al., 2011; Boyer, 2008; Peters & Büchel, 2010; see Bulley et al., 2019 for review of this literature).

One possible explanation for our findings is that some of our measures were not adequately tapping the constructs they were meant to assess. However, we note that the intertemporal choice measures showed meaningful relations either with age or IQ (or both). The EFT procedure closely followed others in the literature, and although the EFT measures did not correlate with age, there were correlations with verbal IQ. It is true that there was one potentially unexpected feature of the EFT data, which was that scores on the EFT task, as well as self-ratings of speed at which the future event came to mind (but not clarity of the imagined event), were actually greater for the most distant time period. This contrasts with Coughlin et al.’s (2014) findings, who found that distance in the future (next week vs. next year) did not affect the amount of episodic detail produced and that children reported bringing events next week to mind quicker than those next year. The time periods we were comparing (tomorrow, next week, and a few months’ time) were different to those used by Coughlin et al., which is likely to explain the differences in findings. It is perhaps not surprising that children found it more difficult to generate rich episodic descriptions for tomorrow as opposed to next week or a few months’ time because the latter two time periods were more extended than a single day and thus encompassed considerably more potential events. Our impression was that for tomorrow events, some children were likely to resort to descriptions of mainly generic events such as those that might happen in a typical school day.

Assuming that our measures really did tap the appropriate constructs, then the findings suggest that the relation between EFT and delay of gratification changes developmentally between middle childhood and adolescence. This may be because EFT itself continues to develop substantially over this period (Abram, Picard, Navarro, & Piolino, 2014; Gott & Lah, 2013), with studies suggesting that such improvements may be linked to improvements in executive functioning (Gott & Lah, 2013), narrative abilities, and self-concept coherence (Coughlin et al., 2014). Although the mechanism underpinning a relation between EFT and delay of gratification is still not fully understood (Rung & Madden, 2018), there is evidence to suggest that engaging in EFT is only beneficial with regard to impulsive choice when sufficient working memory resources are available. Lin and Epstein (2014) found that priming adults to engage in EFT only reduced impulsive choice on a delay discounting task for participants with higher working memory capacity. This suggests that our findings might be explained either because children have lower working memory capacity than adolescents, or because engaging in EFT itself makes greater demands on children than adolescents. Future studies could test these explanations by examining working memory capacities alongside EFT and delay of gratification in children and adolescents.

Conclusion

This study was the first to examine the relations between future time perception, EFT and delay of gratification in children, with delay of gratification measured using both real and hypothetical rewards. We found evidence that children’s performance on the delay discounting task was related to their future time perception, consistent with theoretical claims regarding the link between future time perception and delay of gratification (Kim & Zauber, 2009, 2019). This finding suggests that developmental changes in impulsive choice may be linked to changes in future time perception. However, it is not clear why future time perception was unrelated to performance on a delay of gratification task involving real rewards. Contrary to theoretical approaches linking EFT and impulsive choice, we found no strong evidence that EFT was related to children’s ability to delay gratification, and speculated that this may be due to limited working memory resources in children or how demanding children find it to engage in EFT. As yet, the developmental benefits in middle childhood of engaging in EFT have not been properly elucidated.

References


**Supporting Information**

Additional supporting information may be found in the online version of this article at the publisher’s website:

**Appendix S1.** Supplementary Materials and Analysis.