



# The process of early self-control: an observational study in two- and three-year-olds

Hanna Mulder<sup>1</sup>  · Heleen van Ravenswaaij<sup>2</sup> · Josje Verhagen<sup>1</sup> · Mirjam Moerbeek<sup>3</sup> · Paul P. M. Leseman<sup>1</sup>

Received: 7 December 2017 / Accepted: 1 July 2019 / Published online: 26 July 2019  
© The Author(s) 2019

## Abstract

Early individual differences in self-control are predictive of numerous developmental outcomes, such as physical health and risk-taking behaviours. Therefore, it is important to improve our understanding of how young children manage to exert self-control. This study investigates two- and three-year-old children's behaviours during two self-control tasks and the association between the occurrence of these behaviours and task success. Furthermore, the study examines relations between timing and occurrence of these behaviours and caregiver-rated self-control. Two- and three-year-olds ( $N=62$ ) were given a snack and gift delay of gratification task to measure self-control. The following behaviours were coded second-by-second during the delay: direction of visual attention and the body (directed towards or away from the reward), distracting with the hands (fidgeting), and actively 'withholding' the hands (e.g., holding one hand with the other). To assess caregiver-rated self-control, parents and teachers filled out a selected set of items from the Early Childhood Behaviour Questionnaire inhibitory control scale (Putnam et al. 2006). The percentage of time that children looked away and withheld their hands from the reward positively predicted task success, suggesting that these behaviours are strategic at this young age. Average latency of initiating these behaviours was  $<10$  s for successful children. Teacher-rated (but not parent-rated) self-control related to both the timing and co-occurrence of these behaviours, supporting the ecological validity of the observations. These findings call for future studies to examine further how individual and contextual factors shape the fine-grained dynamics of strategy use in self-control early in life.

**Keywords** Delay of gratification · Early childhood · Self-control · Situational · Attention · Distraction · Strategies

---

Heleen van Ravenswaaij shared first author

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s11409-019-09199-3>) contains supplementary material, which is available to authorized users.

✉ Hanna Mulder  
h.mulder2@uu.nl

Extended author information available on the last page of the article

## Introduction

Self-control refers to the ability to restrain oneself from acting on impulse in situations where this is called for (Hofmann et al. 2009). Self-control is a key aspect of top-down self-regulation (Bridgett et al. 2015; Nigg 2017) and is required for effective delay of gratification – that is, the refusal of an immediate reward in order to reach a long-term goal instead. Already in the first years of life, individual differences in self-control are predictive of a range of adaptive behaviours and outcomes much later that share as common characteristics the valuation of long-term goals over immediate gratification. For example, early childhood self-control is predictive of physical health in adulthood (Schlam et al. 2013), as well as substance dependence, personal finances, and criminal offending (Moffitt et al. 2011).

Although the ability to deploy self-control and associated delay of gratification develops during the preschool years and continues to develop for many years thereafter (Kopp 1982; Manfra et al. 2014), some children as young as 2 years of age have already been observed to successfully delay gratification (Vaughn et al. 1986). Given the importance of early self-control and delay of gratification for a range of developmental outcomes, it is crucial to improve our understanding of *how* such very young children manage to exert self-control (Wilson et al. 2009). Therefore, the current study aims to investigate which behaviours two- and three-year-olds show during the wait period in a delay of gratification situation, and how these behaviours relate to task success. Moreover, we aim to investigate how the occurrence of these behaviours unfolds over the delay time, and how caregiver ratings of children's self-control relate to the occurrence and timing of behaviours during delay.

## Behaviours subserving effective self-control

To date, two main experimental paradigms have been used to assess delay of gratification and associated delay behaviours in young children. First, in preschoolers (i.e., three- to five-year-olds), delay of gratification has most often been assessed using Mischel's classic 'marshmallow' paradigm or a version thereof (see Mischel et al. 1989). In this task, the child is asked to try not to eat a single marshmallow immediately, in order to gain two marshmallows after a period of delay. Second, for toddlers (i.e., two-year-olds), a simpler task has been used in which children are asked to refrain from touching a reward in front of them, such as a wrapped gift, snack, or attractive toy, for a limited period of time (e.g. Kochanska et al. 1997; Vaughn et al. 1986). Crucially, in this toddler version, the prospective larger reward is absent, because providing complex verbal instructions of this type requires linguistic abilities typically not yet in place in most young children. In both the classic delay of gratification task and the toddler version of this task, however, children are directly facing the attractive object in front of them during the delay time. Therefore, both tasks allow detailed coding of behaviours that children show in relation to the reward during delay (e.g., Manfra et al. 2014; Peake et al. 2002).

Over the past few decades, several theories have been posited to explain how self-control – and failure thereof – may come about in delay of gratification situations. These theories, such as the 'hot/cool systems framework' from Metcalfe and Mischel (1999) and the 'process model of self-control' (Duckworth et al. 2016) are based on the idea that a broad range of behaviours related to actively *avoiding* focus on the reward are seen as effective for task success. More specifically, these theories posit that self-control is facilitated by shifting attentional focus away

from the reward and its salient ‘hot’ consummatory properties, and/or making the physical action that is required to reach to the reward more effortful. Next, we discuss empirical findings from the early childhood literature regarding the effectiveness of these types of behaviours for task success.

Research to date shows consistent evidence regarding the effectiveness of self-distractions for successful delay of gratification. Specifically, a range of studies has shown that children who are able to direct visual attention away from the reward are more successful on delay of gratification tasks than children who stay focused on the reward in front of them (e.g., Mischel et al. 1989; Peake et al. 2002, Sethi et al. 2000; Schlam et al. 2013; Vaughn et al. 1986; but see Manfra et al. 2014). Furthermore, talking about something other than the reward tends to be related to task success (Vaughn et al. 1986), and children who are taught to focus on a non-consummatory aspect of the reward are more successful than children who focus on its consummatory properties (‘marshmallows look like clouds’ versus ‘marshmallows taste yummy’, respectively; Mischel and Baker 1975).

In addition to changing their direction of gaze and focus of thought, children may apply *situation selection* or *situation modification* to regulate emotions (Gross 1998, 2015) and, as such, facilitate self-control (Duckworth et al. 2016). In the early childhood literature, these behaviours have generally been described as ‘gross motor strategies’ and ‘fine motor strategies’/‘physical restraint’, respectively. Gross (2015) describes situation selection as “taking actions that make it more or less likely that one will be in a situation that one expects will give rise to desirable or (undesirable) emotions” (p.7). In a study with three- and four-year-olds, children were sometimes observed to walk away to increase the distance between themselves and the reward (Manfra et al. 2014). In fact, in this study, 61% of the children who succeeded in a 3-min delay task showed such ‘moving away’ behaviours, while this was only the case for 6% of the children who failed the task. Thus, young children may already engage in rudimentary forms of situation selection, even within the confines of a delay of gratification task in a testing room.

Duckworth and colleagues (Duckworth et al. 2016) describe situation modification behaviours as “purposefully changing our circumstances to advantage” (p. 40). A frequently used example is the Greek myth of Odysseus who ties himself to the mast of his ship to ensure he could not fall victim to the call of the sirens (c.f., Duckworth et al. 2016; Fujita 2011). More common examples include school-aged children who place their mobile phones out of reach when having to study (Duckworth et al. 2014). In the delay of gratification situation, young children may actively modify their situation in a number of ways even without getting up from their seat. For example, children may sit on their hands so that they cannot reach to the reward directly, put their hands behind their back or place the bell that they can ring to call back the assessor (in a particular variation of the delay protocol) out of reach. Carlson and Beck (2009) showed that occurrence of these types of behaviours positively predicted delay time in three- and four-year olds. To summarize, previous studies have shown that a broad range of behaviours, including directing visual attention away from the reward, talking about something other than the reward, and selecting or modifying the situation to their advantage, relates to success on delay of gratification tasks in young children.

In the current study, we provide a comprehensive assessment of behaviours two- and three-year-olds show during two delay of gratification tasks. Consistent with previous findings from the early childhood literature and theoretical models which center on self-distractions and active avoidance of the reward as key to effective self-control (Duckworth et al. 2016; Metcalfe and Mischel 1999), we coded whether children engaged in attention and verbal

self-distractions, increased the distance between themselves and the reward, and/or made the action through which the reward could be reached more effortful. In keeping with the early childhood literature (e.g., Manfra et al. 2014), we refer to these behaviours as attention, verbal, and motor behaviours from here on.

### Timing of delay behaviours

To gain further insight into how self-control works in young children, it is important to study the timing of behaviours that children show during the delay. Previous studies reveal intriguing and seemingly conflicting findings regarding time effects. A number of studies show that the initiation of self-distractions and active avoidance behaviours may be relatively slow processes in young children. In the study by Manfra et al. (2014), the average latency to initiate these behaviours was 67 and 42 s for pre-schoolers who were successful on the delay task and either showed moving away or self-restraint behaviours, respectively. In a longitudinal study on the timing of self-distraction behaviours during delay at 18, 24, 36, and 48 months, Cole and colleagues (Cole et al. 2011) found that particularly the youngest children took a long time to initiate and fully engage in distraction: about two to 3 min on average at 18 to 36 months (Cole et al. 2011).

Three studies investigating the predictive value of delay of gratification test performance in toddlers and preschoolers for cognitive development in adolescence provide more indirect insight into the timing of these types of behaviours (Friedman et al. 2011, Shoda et al. 1990; Watts et al. 2018). Pooling data from a large number of experiments, Shoda and colleagues (Shoda et al. 1990) found that delay of gratification test performance in four-year-olds predicted adolescent academic outcomes, but the prediction was limited to those test conditions where the reward was present and children were not given any instructions on how to self-distract. In other words, the predictive value of the delay of gratification test for later outcomes occurred when the assessment captured individual differences in children's spontaneous – rather than assessor imposed - self-distraction behaviours. The authors argued that the task – without explicit instructions about strategy use - thus seems to be tapping into key individual differences in children's command of meta-cognitive strategies for self-control. In their theoretical hot/cool systems framework, Metcalfe and Mischel (1999) describe that metacognition is part of the cool system: a slow, reflective system required to overcome stimulus control.

In contrast to these findings, a recent study largely failed to replicate the long-term predictions from delay of gratification test performance at age four to academic outcomes in adolescence after taking into account a relatively stringent set of statistical controls (Watts et al. 2018). This study showed that the prediction of later academic outcomes was mostly limited to the distinction between children who did not manage to wait for the first 20 s versus children who waited 20 s or more. The authors argue that "... [metacognitive] strategies are unlikely to have played much of a role in a child's ability to wait for only 20s." (p. 1172–1175). Thus, this study shows that the behaviours that occur, or that do *not* occur, very early on in the delay time may be most important for characterizing individual differences in self-control in pre-schoolers. Furthermore, in a large longitudinal study on self-control during delay of gratification in younger children, a latent class growth model was investigated to characterize individual differences in development of self-control from age 14, 20, 24, to 36 months (Friedman et al. 2011). A two-class solution with a 'high' and 'low' self-control class fitted the data best. Interestingly, these groups differed only in the probability of touching the reward

immediately (within the first 10 s) versus not at all (>30 s when the task ended), but not in the probability of touching the reward at intermediate delays. Class membership was predictive of executive function development at age 17 years, confirming the validity of the class distinction. Taken together, the latter two studies suggest that in the process of self-control something key happens very early on in the delay time, both in pre-schoolers (Watts et al. 2018), and at younger ages (Friedman et al. 2011) when metacognitive strategies according to some cannot occur yet (too short time frame), although a distinct alternative possibility is that the critical issue is whether children do or do not activate specific behaviours very early on to help them delay. To address this issue further, more fine-grained analyses of the timing of behaviours young children show during delay are needed. Therefore, the current study attempts to investigate the timing of self-distraction and active avoidance behaviours shown by two- and three-year-olds in greater detail than previous studies have done, using second-by-second behavioural coding and multilevel analyses to investigate time effects on occurrence of behaviours. Given that there are, to the best of our knowledge, no previous studies which used a similar analytic approach to the study of timing of delay behaviours, we conducted an explorative and broad investigation, focusing on the full delay time (rather than a limited section of time) and including both linear and nonlinear time effects in the multilevel models.

### **Parent- and teacher-rated self-control**

Finally, we investigate how self-control as rated by caregivers is associated with the occurrence and timing of behaviours shown during delay over time, to improve our understanding of the relevance of the observed delay behaviours beyond the specific test situation. Results from a study on older children show that caregiver ratings of self-control are associated with children's behaviour during delay of gratification (Wilson et al. 2009). Specifically, Wilson et al. (2009) constructed a 'difficulty with delay' variable from 8 to 11-years olds' behaviours such as fidgeting, inquiry about the reward, focus on the reward, and body tension during delay of gratification. Children who were rated higher on self-control in a questionnaire had significantly lower scores on this difficulty with delay measure (which is indicative of absence of self-distraction and avoidance behaviours). Further, adult studies show that individuals scoring high on self-control questionnaires may naturally recruit strategies to select and modify their situation to avoid temptations, as they pro-actively avoid situations that directly challenge their self-control (Hofmann et al. 2012; for reviews see Duckworth et al. 2016; Fujita 2011). For example, they reported choosing situations and circumstances that minimized the risk of having to directly face temptations (e.g., through choosing to work in a quiet rather than distracting environment when a task required concentration) (Ent et al. 2015).

### **Current study**

The overall aim of the current study is to improve our understanding of the occurrence, timing, and function of a range of behaviours two- and three-year-old show during delay. To this end, we coded the visual attention, verbal, and motor behaviours during delay of two- and three-year-old children performing two 1-min delay tasks. We address four specific aims. First, we provide a description of the occurrence of each of these behaviours. Based on previous literature, we expected that children would show a broad range of distracting and avoiding behaviours during delay, such as directing visual attention away from the reward, turning away

from the reward altogether, and the child changing his or her physical circumstances to advantage, through for example placing their hands under the table (like situation modification in the adult literature on self-control) (Carlson and Beck 2009; Manfra et al. 2014; Mischel et al. 1989; Sethi et al. 2000; Vaughn et al. 1986). Second, we investigated how occurrence of each of these behaviours predicted task success. We expected that self-distraction and active avoidance behaviours would relate positively to task success, based on earlier studies with young children (Carlson and Beck 2009; Manfra et al. 2014; Mischel et al. 1989; Sethi et al. 2000; Vaughn et al. 1986). Third, we studied the timing of behaviours children showed during delay. We stopped short of developing specific hypotheses regarding the timing of delay behaviours, given previous seemingly contrasting findings (Cole et al. 2011; Friedman et al. 2011; Manfra et al. 2014; Shoda et al. 1990; Watts et al. 2018) and because not much is known yet about time effects of behaviours young children show during delay in general. Fourth, we investigated how parent- and teacher-ratings of children's self-control related to the occurrence of these behaviours across time. Based on the literature on adults, we predicted that, also at this much younger age, children with higher caregiver ratings of self-control would be more strongly avoidant of the reward and would initiate avoidant behaviours very rapidly (Duckworth et al. 2016; Ent et al. 2015; Fujita 2011; Hofmann et al. 2012).

## Method

### Participants

Data for the current study were collected as part of a study aimed at developing an assessment battery of language and executive function measures for two- and three-year-olds for use in a cohort study (see Mulder et al. 2014). From this sample, we selected only children for whom a video of at least one of the two delay of gratification tasks (see below) was available. This resulted in a sample of 62 children with a mean age of 34.8 months ( $SD = 5.7$ ; range 23.5 to 42.7) and 29 (46.8%) girls. The majority of children (38 out of 55 with available parent questionnaire data, 69.1%) were from families in which one or both parent(s) had completed at least higher education (i.e., college or university). For 7 children, this information was missing. For all children for whom information on home language was available, Dutch was spoken at home (56 available reports, 6 missing).

### Procedure

The delay of gratification tasks were intermixed with a series of language and executive function tasks not reported on in this paper, and administered in a fixed order. First, children were given tasks assessing selective attention and language abilities using a laptop, that are not reported on in this study. Next, the snack delay task was administered. The gift delay task was given approximately 20 min later, after a series of working memory tasks which are not part of the current report. All children also took part in a second assessment session about 2 weeks later, which included further executive function and language measures, none of which are reported on in the current study. All tasks were administered by extensively trained research-assistants (RAs), who were enrolled in a master's degree programme Clinical Child, Family, and Education Studies. A quiet

room with as few stimuli as possible in the children's daily environment (i.e., home [ $n = 8$ ] or day-care centre [ $n = 54$ ]) was used for testing. Parents were informed about the study through an information letter and brochure, and signed for their child's participation.

The current study aimed to develop measures for a large field research with children of varying ages. Therefore, we used a version of the snack and gift delay task with a 60-s delay in part of the sample of three-year-olds, and a 150-s delay in another part of the sample of three-year-olds, to see if more variance in pass/fail scores could be obtained when working with a longer delay in this age group. For consistency, however, only the first 60 s of both versions were coded for the purpose of the current study, and data from all children were pooled in the analyses. Also, for the children taking the longer version of the task, another task not reported on in this article was given before the gift delay task, in which the assessor was pretending to wrap the gift while the child was instructed not to peek (Kochanska et al. 1996) (14/40, 35% of three-year-olds). As the tests were administered at children's homes or day-care centre, the size and height of the test table and chair were not standardized (24/62, 39%, were seated at a low child-height table on a small child's seat, all others were seated at an adult-size table), and a caregiver was present in the room during test administration occasionally (17/60, 28%, for snack delay, 13/55, 24%, for gift delay). The potential influence of these variables (i.e., test version [short version without the additional part where the assessor was pretending to wrap the gift versus longer version with the additional part], table/chair height, and presence of caregiver) was studied in a series of preliminary analyses, as further described in the Analysis section.

## Measures and coding

**Observed self-control** Two delay of gratification tasks, a snack and a gift delay task, were used to measure self-control. These were adapted for field research from the effortful control task battery of Kochanska and colleagues (Kochanska et al. 1996; Kochanska et al. 2000). In the snack delay task, children were shown an open box of raisins and asked to try not to touch the box of raisins until the RA had finished another task. Similarly, in the gift delay task children were shown an attractively wrapped gift with a bow and asked to try not to touch the gift until the RA had completed another task. RAs were trained on giving the instructions for both tasks as a friendly request (e.g., by saying 'try not to touch' in a friendly voice). Directly following the instructions, the reward was placed on the table at a distance of 25 cm from the child. In both tasks, the assessor pretended to be writing while sitting at a short distance behind the child during the 60-s delay time. At the end of the delay time, children always received positive feedback; they were either praised for waiting, or in case of 'failure', asked whether the raisins tasted good or complimented on the nice gift they got. Before the task was administered, caregivers (i.e., parents or teachers, depending on the setting that the test was administered in) were explained the purpose of the task and asked to either leave the room for the short duration of the task, or sit down quietly elsewhere in the same room and pretend to be reading, so as not to interact with the child during the delay time. Children's performance on each of the delay of gratification tasks was scored as *fail* if they touched the reward at least once before the 60 s had passed and as *pass* if they completed the task without touching the reward for 60 s.

**Behaviours shown during delay** Children's behaviours during the 60-s delay time were coded from video by the second author for both tasks. On all five domains (visual attention,

verbal, and hand, head, and body behaviours) directing attention towards the reward was coded as focusing, whereas directing attention away from the reward (i.e., self-distraction) and active avoidance was coded as distracting and withholding, respectively. Withholding can be seen as a special way of exerting self-control, which includes behaviours that children can use to actively stop themselves from touching the reward, such as placing their hands under the table. When a child showed different behaviours with their right and left hand, the hand regarded as closest to touching the reward was coded. Examples of codes within each of these categories are presented in Table 1; the full coding scheme is shown in Appendix A. The shared first author (HvR) coded all domains separately on one-second intervals using Mediacoder (Bos and Steenbeek 2009). That is, the video was paused every second and codes were based on the video still for all behavioural domains except the verbal domain. For the verbal domain, transcripts were generated first, and each utterance was coded separately with start and stop times. This approach resulted in five codes per second and 300 codes in total per video. Note that we coded focusing, distracting, and withholding within each of the domains to ensure comprehensive coding of the observed behaviours. However, only specific categories representing distracting and avoidance behaviours that were hypothesized to support self-control and that passed the inter-coder reliability test, discussed below, were used in further analyses; these specific variables are marked with an (\*) in the Appendix and in Table 1.

In addition to the raw data, two measures for each of the five domains were derived for analytic purposes. First, the percentage of intervals that each behaviour occurred was computed for each child, across all time intervals until the child touched the object for the first time (or until a maximum of 60 s was reached for children who passed the task). For children who did not show a particular behaviour at all, the percentage occurrence variable was zero. Second, a latency variable was constructed, which refers to the time it took children to show each type of delay behaviour for the first time. This variable was constructed only for children who showed a particular behaviour at least once. As such, children who did not show a particular behaviour had a missing value on the latency variable for that behaviour.

**Table 1** Inter-rater reliability for each domain with examples of codes

Domain	Kappa (95% CI)	Subcode	Example of behaviour	Sub Kappa <sup>a</sup> (95%CI)
Task pass/ fail	.96 (.94–.98)	n/a	Touches reward	–
Visual attention	.76 (.71–.80)	Focusing	Eyes focus on reward	.79 (.75–.83)
		Distracting*	Eyes away from reward	.77 (.72–.81)
		Withholding	Eyes are closed	.15 (–.05–.35)
Verbal	.36 (.26–.46)	Focusing	Talks about reward	.00 (.00–.00)
		Distracting	Reward-unrelated talk/ mumbling	.22 (.11–.33)
		Withholding	Repeats the rule	.83 (.70–.96)
		Does not speak	–	.35 (.25–.45)
Hands	.81 (.78–.85)	Focusing	Almost touching the gift	.80 (.76–.84)
		Distracting*	Touches mouth/ fidgeting	.68 (.61–.76)
		Withholding*	Hands underneath the table	.91 (.88–.93)
Head	.65 (.60–.70)	Focusing	Facing the reward	–
		Distracting	Not facing the reward	–
Body	.66 (.61–.71)	Focusing	Torso turned to reward	–
		Distracting*	Torso turned away from reward	–

All values based on behaviours of successful and unsuccessful children combined, except for hands category, which was based on successful children only. <sup>a</sup> Only for domains with >2 categories. \*Behavioural categories that were used in the analyses

**Reliability** Fifteen percent of the available videos were randomly selected and coded by a another coder to establish inter-rater reliability. We computed Kappa to evaluate reliability, for each domain as a whole (visual attention, verbal, etc.), and for each specific behaviour within each domain (presence vs. absence of this behaviour, only for domains with at least two behavioural codes) (see Table 1). Based on Landis and Koch (1977), a Kappa of .61 or higher was considered acceptable.

Inter-rater reliability was acceptable for the visual attention, hands, and body domains, somewhat lower for the head domain (95% CI just overlapped with .60) and very low for the verbal domain. As it was somewhat challenging to obtain agreement between coders on the direction of the head, and direction of the head was relatively strongly related to direction of the eyes (Kendall's tau = .43;  $p < .001$  for snack delay; Kendall's tau = .57;  $p < .001$  for gift delay), which we considered the more important variable, the former was dropped from the analyses.

Closer inspection of the reliability data for the verbal domain revealed that the low reliability was likely due to a combination of factors: an extremely unequal cross-cells distribution with many empty cells (e.g., Kottner and Dassen 2008), and difficulty to distinguish distracting from 'no speech' successfully. The latter issue was probably due to the fact that we also coded 'mumbling' as distraction, and establishing whether a child was mumbling quietly or making mouth movements without speaking was sometimes difficult. For these reasons, the verbal domain was dropped from the analyses.

Inter-rater reliability at the level of each specific behaviour showed that, within the visual attention domain, withholding was not reliable. As occurrence of this behaviour was very infrequent (1.3% and 1.5% of intervals on average on snack and gift delay, respectively), it was combined with visual attention distraction in the analyses. Finally, as some of our analyses involved concomitant behaviours (see further under Analytic section below), inter-rater reliability of the co-occurrence of behaviours was investigated. Co-occurrence of the visual attention, hands, and body domains was reliable between coders (Kappa = .65; 95% CI = .61–.69).

**Data screening** Nine videos (i.e., two of the snack and seven of the gift delay task) had to be excluded due to low technical video quality or task administration errors (e.g., caregivers interacting with the child during the task). In total, video data of 60/62 (93.5%) children were available for the snack delay task, and video data of 55/62 (87.1%) children were available for the gift delay task.

**Parent- and teacher-rated self-control** Parents and teachers were asked to rate children's self-control on a number of selected items from the inhibitory control subscale of the Early Childhood Behaviour Questionnaire (ECBQ; Putnam et al. 2006). This subscale assesses children's ability to moderate behaviour or refrain from acting in situations where this is called for. The ECBQ is suited to measure temperament in children aged 18 to 36 months, and an age-adjusted version is available for children aged 3 years and older, the Child Behaviour Questionnaire (CBQ; Rothbart et al. 2001). However, our study included both two- and three-year-olds and the same scale needed to be given to all children in the sample for analytic purposes. Therefore, we decided to use the ECBQ for all children. In addition, for the purpose of the large field study, to limit the time required from parents and teachers to fill out the questionnaire, we used only a very limited set of ECBQ items that we had previously selected based on an earlier pilot as described in

(Mulder et al. 2014).<sup>1</sup> Five items were given to parents, and three items were given to teachers. An example item given to parents is: 'When asked not to, how often did you child touch an attractive item anyway?'. An example item given to both parents and teachers is: 'When told no, how often did this child ignore your warning?'. For each item, parents and day-care teachers were asked to indicate how often the child displays the respective behaviour on a scale from 1 (*never*) to 7 (*always*). A total of 55 parents (55/62, 87%) and 50 day-care teachers (50/62, 81%) completed the items.

Internal consistency (Cronbach's alpha) of our short version was .68 for parents and .94 for teachers in the current study. Moreover, in the current study, we administered the standard version of the inhibitory control scale of the ECBQ to parents and teachers of two-year-olds (12 items), and the standard short version of the inhibitory control scale of the CBQ to parents and teachers of three-year-olds (6 items). The association between inhibitory control on the standard ECBQ form and our short version was .84 ( $p < .001$ ) for parents and .87 ( $p < .001$ ) for teachers of two-year olds. The correlation between inhibitory control on the standard short CBQ form and our short version was lower at .42 ( $p = .013$ ) for parents and .61 ( $p < .001$ ) for teachers, likely due to the broader scope of the CBQ. Specifically, unlike the ECBQ, the CBQ includes items related to the ability to follow instructions and plan ahead (such as 'Prepares for trips and outings by planning things s/he will need').

## Data analyses

To investigate which behaviours were predictive of task success (pass/fail on each task), a series of analyses were conducted. First, in a number of preliminary analyses, we investigated whether a set of variables (i.e., age, gender, chair height, the presence of a caregiver in the room, and test version) influenced the results and, hence, should be controlled. These variables were studied in relation to task success using logistic regression. Subsequently, we studied whether these variables predicted the percentage of time intervals that children showed each behaviour during delay, using linear regression. The variables that related to both task performance and/or behaviours during delay were included in the main analyses. Second, in the main analyses, we entered the behaviours children showed during delay in a logistic regression analysis with task performance (pass/fail) as the outcome variable to investigate which behaviours were predictive of task success. Bootstrapping was used in these analyses, because of the non-normal distribution of some of the variables (e.g., the percentage of time a behaviour occurred) and the small sample (Efron and Tibshirani 1993).

To investigate the effects of time and caregiver ratings of self-control on occurrence of behaviours during delay, a series of multilevel multinomial regression analyses were run in HLM (version 7.03). Multinomial regression is an extension of binary logistic regression, and allows the dependent variable to have more than two unordered categories (Hedeker 2003). We opted for this type of analyses because two behaviours were significantly related to delay task performance: visual attention distraction and hands withholding. To gain comprehensive

<sup>1</sup> In the field study, these parent and teacher items were used as indicators to a latent inhibitory control construct. The correlation between this construct and a latent delay of gratification construct with performance on a snack and gift delay task as indicators, was .53 in two-year-olds (Mulder et al. 2014).

insight into the unfolding of these behaviours over time, their joint occurrence was modelled as a single dependent variable with  $2 \times 2$  categories. Thus, the dependent variable had the following four categories: 1) visual attention focusing and hands not withholding (this is the least controlled combination of behaviours), 2) visual attention focusing and hands withholding, 3) visual attention distracting and hands not withholding, and 4) visual attention distracting and hands withholding (this is the most controlled combination of behaviours). The first category was set as the reference category in the analyses. The logit probability of the occurrence of each of the other three categories was modelled relative to the reference category.

In the multilevel model, time was the first level predictor and caregiver-rated self-control was the second (child-) level predictor. The cross-level interactions between time and caregiver ratings were also included. The following models were run: 1) linear and quadratic time effects models to investigate both linear and nonlinear time effects; 2) time and caregiver ratings effects models with cross-level interactions and age as covariate. The models were fitted with the Penalized Quasi-Likelihood estimator that is available in HLM. As no relative fit indices are available for model comparison with this estimation method, we investigated statistical significance of each of the polynomial time effects in step 1 to determine whether they should be included in the model or not. All independent variables were added to the model grand mean centred.

Note that, although the number of observation points at the first level can be regarded as rather substantial for an observational multilevel study (Hox et al. 2018), the number of cases at the second level was relatively small. We therefore closely inspected our models for the stability of findings. All models ran without difficulty and the maximum number of iterations required was 27. Bonferroni correction was applied to adjust for a potential inflation of the Type I error rate. There were three tests of the statistical significance of the independent variables (time, parent and teacher rated self-control) for each task, that is, one test for the occurrence of each behavioural category in comparison to the reference category. Therefore, alpha was set to  $.05 / 3 = .017$ . Results with robust standard errors are presented.

## Results

### Descriptives of task performance

**Task success** In total, 70% (42/60) of the children managed to delay on request during the snack delay task and 73% (40/55) of the children managed to delay during the gift delay task. Task success was relatively consistent between test versions: children who touched the snack had an 80% chance of also touching the gift (i.e., 12 out of 15 children who touched the snack and had data on both tasks, also touched the gift).

**Latency to delay** For children who touched the snack, average latency of delay was 6.2 s ( $SD = 10.6$ ; range 1 to 38;  $n = 18$ ). The distribution of delay latencies was strongly skewed within the subsample of children who touched the snack: 83% (15/18) touched the snack within the first 10 s, and the majority of those ( $n = 13$ ) touched the snack within the first 3 s. For children who touched the gift, a similar pattern emerged. Average latency to delay was 10.1 s ( $SD = 19.0$ ; range = 1 to 59;  $n = 15$ ), 87% (13/15) children touched the gift within the first 10 s, and the majority of those ( $n = 10$ ) touched the gift within the first 3 sec. In both tasks,

a number of children touched the reward at the very first second ( $n = 7$  for snack delay,  $n = 4$  for gift delay). These were mostly two-year-olds ( $n = 5$  for snack delay,  $n = 3$  for gift delay).

### Descriptive statistics of behaviours shown during delay

Descriptive statistics of the behaviours children showed during delay are given in Table 2, separately for children who passed and failed the task. Table 3 shows intercorrelations between behaviours for both tasks, teacher and parent reports of children's self-control, and age. Raw behavioural data are shown in Figures 1 and 2.

Distracting visual attention from the reward occurred in all children who passed the snack delay task at least once during the delay time (100%), followed by body distracting (91%), hands withholding (85%), and hands distracting (54%). Occurrence of each of these behaviours was generally lower in children who failed the task, that is,  $\leq 22\%$  for each of the behaviours. The latency data show that initiation of visual attention distraction and hands withholding was rapid in successful children, at an average of 4 and 7 s, respectively, while initiation of hands and body distraction behaviours occurred somewhat later, i.e., between 12 and 18 s on average. The raw data plotted in Fig. 1 show that hands withholding behaviours already occurred at the very first second of the delay time in about half of all successful children, and about a third of successful children showed visual attention distraction immediately at the outset of delay.

A similar pattern of findings emerged for the gift delay task. In successful children, visual attention, hands, and body distracting behaviours and hands withholding occurred at least once in  $\geq 54\%$  of cases, compared to  $\leq 40\%$  in unsuccessful children. Initiation of visual attention distraction and hands withholding was rapid in successful children, on average within about the first 5–6 s, while initiation of hands and body distraction occurred somewhat later, after the first 14–22 s of the delay time had passed on average. As shown in Fig. 2, about two-thirds of the successful children showed hands withholding already during the first second, while a quarter showed visual attention distraction at this time.

### Preliminary analyses: Age, gender, chair height, presence of caregiver, and test version

A logistic regression analysis in which snack delay performance was regressed on age, gender, chair height, and presence of a caregiver was significant ( $\chi^2(4) = 20.4$ ;  $p < .001$ ). Age was the only significant predictor, as older children had lower odds of failing the task ( $B = -0.24$ ;  $SE = 0.12$ ;  $p < .001$ ; 95% CI of  $B = -0.54$  to  $-0.14$ ). For gift delay, a logistic regression with age, gender, chair height, presence of caregiver, and task version as predictors was significant ( $\chi^2(4) = 13.8$ ;  $p = .017$ ), with age again as the only significant predictor ( $B = -0.17$ ;  $SE = .26$ ;  $p = .012$ ; 95% CI of  $B = -0.57$  to  $-0.02$ ).

Next, a linear regression was conducted in which behaviours shown during the snack delay task were regressed on age, gender, chair height, and presence of a caregiver. The regression model for the snack delay task for distracting visual attention was significant ( $R^2 = .38$ ;  $p < .001$ ). Age and gender were significant predictors of the percentage of time children showed visual distraction. Older children were more likely to show visual distraction than younger children ( $B = 2.5$ ;  $SE = 0.7$ ;  $p = .003$ ; 95% CI of  $B = 1.0$  to  $3.8$ ), as were girls as compared to boys ( $B = 20.7$ ;  $SE = 7.5$ ;  $p = .013$ ; 95% CI of  $B = 6.2$  to  $35.2$ ). The regression models for the percentage of time children showed withholding or distracting behaviours with their hands and

**Table 2** Descriptive statistics for each of the behaviours during snack and gift delay by task success

	N children who showed behaviour at least once (%)				Mean latency to first (SD; range)	Mean % of time intervals (SD; range)	Fail (w/o t = 1 touchers) (n = 11)
<b>Snack delay</b>	Pass (n = 42)	Fail (n = 18)	Pass <sup>a</sup>	Fail <sup>a</sup>	Pass (n = 42)	Fail (w/o t = 1 touchers) (n = 11)	
Visual attention distracting	42 (100%)	3 (17%)	3.5 (2.5; 1–12)	2.7 (0.6; 2–3)	79 (13; 40–98)	17 (33; 0–89)	
Hands distracting <sup>b</sup>	22 (54%)	2 (11%)	18.1 (16.5; 1–58)	7.0 (7.1; 2–12)	16 (26; 0–93)	9 (27; 0–89)	
Hands withholding <sup>b</sup>	35 (85%)	3 (17%)	7.3 (13.0; 1–57)	1 (all cases)	65 (39; 0–100)	13 (31; 0–94)	
Body distracting	38 (91%)	4 (22%)	11.9 (13.9; 1–60)	3.0 (2.4; 1–6)	34 (37; 0–100)	32 (45; 0–100)	
<b>Gift delay</b>	Pass (n = 40)	Fail (n = 15)	Pass <sup>a</sup>	Fail <sup>a</sup>	Pass (n = 40)	Fail (w/o t = 1 touchers) (n = 11)	
Visual attention distracting	40 (100%)	6 (40%)	4.7 (5.5; 1–29)	2.7 (2.7; 1–8)	69 (20; 27–97)	31 (37; 0–100)	
Hands distracting <sup>b</sup>	21 (54%)	2 (13%)	22.3 (14.5; 1–56)	1 (both cases)	17 (26; 0–98)	16 (35; 0–88)	
Hands withholding <sup>b</sup>	35 (90%)	2 (13%)	5.6 (12.0; 1–58)	1.5 (0.7; 1–2)	61 (40; 0–100)	16 (35; 0–88)	
Body distracting	34 (85%)	3 (20%)	13.9 (14.2; 1–60)	4 (5.2; 1–10)	27 (32; 0–98)	15 (32; 0–100)	

<sup>a</sup> Only for children who showed behaviour at least once. <sup>b</sup> One successful child's hands data is missing due to a suboptimal position of the camera

**Table 3** Correlations (Kendall’s tau) between percentage of time intervals that each behaviour occurred during delay, caregiver-rated self-control, and age

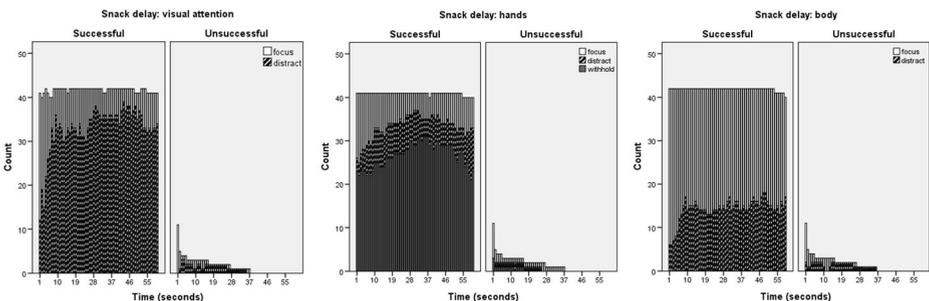
	Snack delay task				Gift delay task				Caregiver rated self-control	Other	
	1.	2.	3.	4.	5.	6.	7.	8.	9.	.10	11.
<b>Snack delay task</b>											
1. Visual attention distracting	–	.20	.26*	.19	.39***	–.12	.19	.08	–.05	.26*	.31*
2. Hands distracting		–	–.22*	.15	.16	.19	–.16	.20	–.12	.12	–.02
3. Hands withholding			–	–.09	.09	–.15	.35**	–.21	.26*	.12	.25*
4. Body distracting				–	.13	.07	–.04	.47***	–.06	.22	.12
<b>Gift delay task</b>											
5. Visual attention distracting					–	.17	.22*	.19	.09	.22	.15
6. Hands distracting						–	–.23*	.26*	–.06	–.05	.01
7. Hands withholding							–	–.02	.29*	.31*	.24*
8. Body distracting								–	.04	.08	.12
<b>Caregiver rated self-control</b>											
9. Parent ECBQ									–	.21	.12
10. Teacher ECBQ										–	.35**
<b>Other</b>											
11. Age											–

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

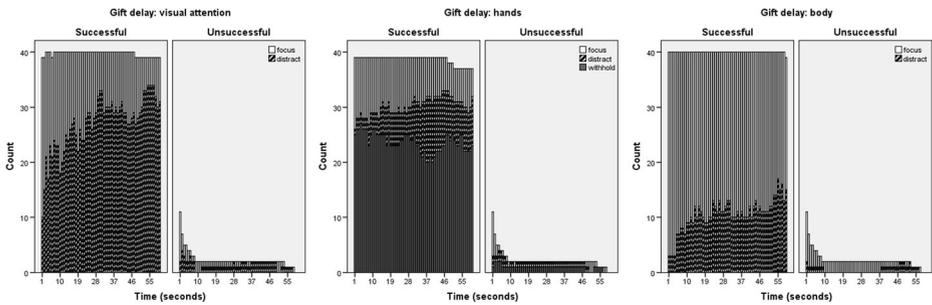
body distracting during the snack delay task were not significant ( $R^2 = .11$ ;  $p = .251$ ;  $R^2 = .03$ ;  $p = .868$ ;  $R^2 = .08$ ;  $p = .459$ , respectively).

For the gift delay task, regression models were run with age, gender, chair height, presence of caregiver, and test version as predictors of behaviours shown during delay. These models were not significant for visual attention, hands, and body distracting ( $R^2 = .16$ ;  $p = .174$ ,  $R^2 = .12$ ;  $p = .356$ ,  $R^2 = .18$ ;  $p = .128$ , respectively). The regression model for hands withholding was trend-level significant ( $R^2 = .20$ ;  $p = .087$ ), due to a significant age effect: older children were more likely to show hands withholding behaviours ( $B = 4.1$ ;  $SE = 1.4$ ;  $p = .006$ ; 95% CI = 1.04 to 6.8).

To summarize, only age was significantly related to task success and some of the behaviours shown during delay. Therefore, age was retained as a covariate in the main analyses for both tasks.



**Fig. 1** Number of cases showing each of the behaviours at each time interval during snack delay, split by task success ( $N = 42$  successful children, one successful child’s hands data is missing due to a suboptimal position of the camera;  $N = 18$  unsuccessful children). Observational data were coded until the moment the child touched the reward for the first time. In the group of unsuccessful children, many children touched the reward relatively early on during the delay time. Hence, the number of cases with observation data rapidly drops over time in this group. Data of  $N = 7$  children are not shown at all as they failed at  $t = 1$



**Fig. 2** Number of cases showing each of the behaviours at each time interval during gift delay, split by task success ( $N = 40$  successful children, one successful child's hands data is missing due to a suboptimal position of the camera;  $N = 15$  unsuccessful children). Observational data were coded until the moment the child touched the reward for the first time. In the group of unsuccessful children, many children touched the reward relatively early on during the delay time. Hence, the number of cases with observation data rapidly drops over time in this group. Data of  $N = 4$  children are not shown at all as they failed at  $t = 1$

### Main analyses (I): How do behaviours shown during delay predict task success?

In a subsequent series of analyses, we entered age and the percentage of time that each behaviour occurred as predictors of task success. The pattern of results was highly consistent between tasks, as shown in Table 4: the percentage of time children spent distracting visual attention away from the object and withholding with the hands were predictive of task success for snack and gift delay, while hands and body distracting behaviours were not.

### Main analyses (II): Time effects on co-occurrence of visual attention distraction and hands withholding behaviours

Next, we studied how the occurrence of visual attention distraction and hands withholding behaviours changed over time. To this end, these variables were combined and entered as dependent variable in multilevel multinomial regression. Linear and quadratic effects of time on the logit probability of each behavioural category (hands withholding only category, visual attention distraction only category, and hands withholding + visual attention distraction combined category) were estimated relative to the reference category (no hands withholding, no visual attention distraction).<sup>2</sup> Results are shown in Table 5 in the online Appendix for successful and unsuccessful children combined. Tables 6 through 9 in the online Appendix show the same analyses for children with questionnaire data only for successful and unsuccessful children combined (Table 6 and 7) and successful children only (Table 8 and 9).

Table 5 shows that there were linear and quadratic effects of time on the logit probability of the occurrence of some of the behavioural categories relative to the reference category for the snack delay task, and only linear effects of time for the gift delay task. Tables 5 to 9 show that the pattern of results in terms of direction of effects was consistent between analyses, but that the level of statistical significance varied by sample selection. We therefore focus our interpretation of these results on the general trends shown in Figs. 3a and b, rather than statistical significance. These Figures show the probability of occurrence of each behavioural category, including the reference category. Probabilities were computed from the model estimates shown in Table 5.

**Table 4** Logistic regression results predicting task success by percentage of time intervals that each behaviour occurred during delay

	B (95% CI)	SE	OR
<b>Snack delay</b>			
Model 1: $N = 53$ ; $\chi^2(2) = 34.65^{***}$			
Visual attention distracting	-0.08 (-1.12 to -0.05)**	0.53	.93
Model 2: $N = 52$ ; $\chi^2(2) = 9.77^{**}$			
Hands distract	-0.01 (-8.16 to 0.04)	5.79	.99
Model 3: $N = 52$ ; $\chi^2(2) = 18.27^{***}$			
Hands withhold	-0.03 (-3.67 to -0.01)*	1.18	.97
Model 4: $N = 53$ ; $\chi^2(2) = 9.70^{**}$			
Body distract	0.01 (-0.04 to 0.03)	1.61	1.01
<b>Gift delay</b>			
Model 1: $N = 51$ ; $\chi^2(2) = 17.48^{***}$			
Visual attention distracting	-0.05 (-0.13 to -0.01)**	0.13	.96
Model 2: $N = 50$ ; $\chi^2(2) = 7.71^*$			
Hands distract	-0.01 (-8.07 to 0.02)	2.46	.99
Model 3: $N = 50$ ; $\chi^2(2) = 13.89^{**}$			
Hands withhold	-0.03 (-23.53 to -0.01)*	9.65	.97
Model 4: $N = 51$ ; $\chi^2(2) = 7.76^*$			
Body distract	-0.01 (-9.12 to 0.02)	20.51	.99

Age is included as a covariate in each model. Children who failed the task at  $t = 1$  and thus had no behavioural observation data, were excluded from the analyses. B, 95% CI,  $p$ -values and SE based on bootstrapped results with 1000 resamples. The dependent variable is coded as 0 = pass, 1 = fail. \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

The figures show a relatively mixed pattern of results between the snack and gift delay task. For the snack delay task, shown in Fig. 3a, the probability of occurrence of the reference category is highest at the outset of the delay time and then decreases over time, while the probability of occurrence of visual attention distraction alone very slightly increases at about 5 s and subsequently decreases and then increases again over time. Probability of occurrence of visual attention distraction combined with hands withholding increases steadily over time, peaks around 30 s, and remains the most probable behaviour to occur after that.

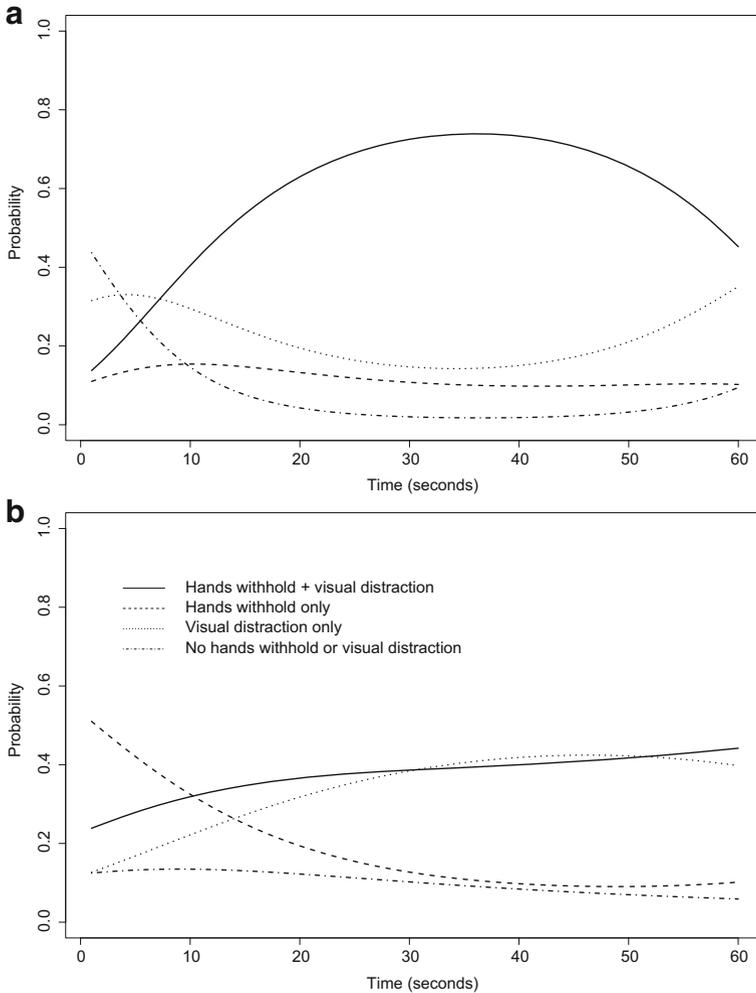
For the gift delay task, shown in Fig. 3b, time effects were generally less pronounced than for the snack delay task. The probability of occurrence of the reference category is relatively stable and low throughout. The outset of the delay time is characterized by a relatively high probability of occurrence of hands withholding only, which subsequently decreases over time. After the first 10–20 s, the combined category and visual attention distraction only prevail.

To summarize, one relatively consistent finding across the two tasks was that probability of occurrence of visual attention distraction only and visual attention distraction combined with hands withholding prevail after about 10–30 s.

### Main analyses (III): Caregiver-rated self-control as predictor of occurrence of visual attention distraction and hands withholding behaviours

Caregiver ratings of self-control and cross-level interactions between caregiver ratings and time were added as independent variables to the multilevel multinomial regression models.

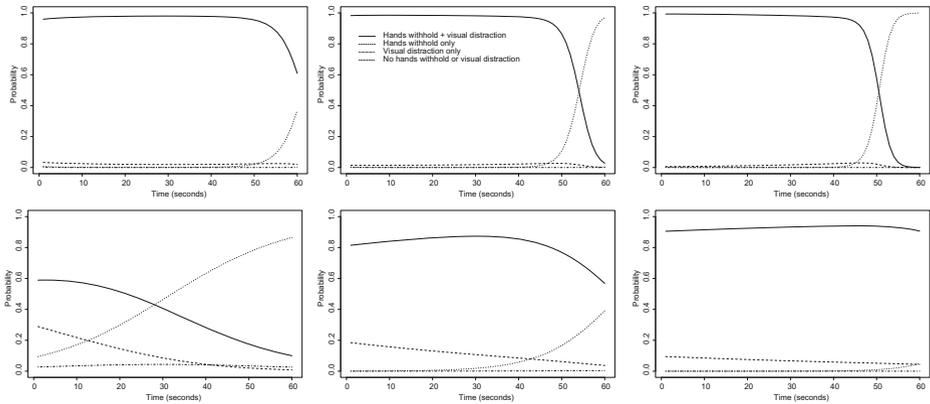
<sup>0</sup> The inter-rater reliability of this particular combination of variables was .78 (95%CI of kappa .75–.82).



**Figs. 3 a and b** Estimated probability of behaviours shown during snack (A, top panel) and gift (B, bottom panel) delay by time

Data from successful and unsuccessful children were combined in these analyses. However, as it is uncertain how task success may influence the pattern of associations between caregiver-rated self-control and behaviours shown during delay, models were also run separately for successful children only. Since the results were consistent between these two sets of analyses, we focus the results description on the full sample. Results are shown in Fig. 4 (full sample) and Tables 6–7 (full sample) and 8–9 (successful children only) in the online Appendix.

There were no significant effects of parent-rated self-control on the logit probability of the occurrence of the visual attention distraction, hands withholding, and combined categories relative to the reference category on either the snack or gift delay task. There was a significant interaction effect of teacher-rated self-control by time on the visual attention distraction category on the snack delay task. Figure 4 shows the estimated probability of occurrence for each of the behavioural categories by time and teacher-rated self-control. For visualization



**Fig. 4** Estimated probability of behaviours shown during snack (top panel) and gift (bottom panel) delay by teacher-rated self-control and time. The estimated probability was modelled for children with low teacher-rated self-control ( $< -2SD$  of the mean, left panels), average teacher-rated self-control (middle panels) and high teacher-rated self-control ( $>2SD$  of the mean, right panels) separately for visualization purposes. The legend in the middle top panel holds for all panels

purposes, the latter variable was split into low ( $2SD$  below the mean, left panel), average (middle panel), and high ( $2SD$  above the mean, right panel) groups. Although the general pattern of occurrence was similar for these three groups, there was a subtle shift in timing: children with average to high teacher-rated self-control shifted from the combined hands withholding + visual attention distraction category to the visual attention distraction alone category more quickly than children with low teacher-rated self-control. Finally, there was a significant positive main effect of teacher-rated self-control on the hands withholding + visual attention distraction combined category and the hands withholding alone category on the gift delay task. Figure 4 shows that children with higher teacher-rated self-control showed more hands withholding behaviour throughout the delay time compared to children with lower teacher-rated self-control.

## Discussion

The current study aimed to improve our understanding of the nature of early self-control through conducting second-by-second observations of behaviours shown by two- and three-year-olds during two 60-s delay of gratification tasks. As predicted based on earlier work on self-control in young children (Carlson and Beck 2009; Manfra et al. 2014; Mischel et al. 1989; Sethi et al. 2000; Vaughn et al. 1986), children engaged in self-distractions and active avoidance behaviours, and occurrence of these behaviours was positively predictive of task success. Successful children initiated these behaviours very early, within the first 10 sec of the delay time on average. Moreover, occurrence and timing of these behaviours related to teacher ratings of self-control, speaking to the ecological validity of these observations.

The current study replicates previous work highlighting the importance of visual attention distraction for success on delay of gratification tasks (Mischel et al. 1989; Peake et al. 2002; Sethi et al. 2000; Vaughn et al. 1986). Furthermore, the current study extends previous findings by highlighting the relevance of ‘withholding’ hands from the reward in an active manner.

With respect to the latter type of behaviour, children frequently tended to block the direct route through which they could touch the reward. For example, they were observed to hold their hands under the table, to cross their arms, or to hold one hand tightly with the other. In reference to the literature on adults and older children, this may be seen as a form of ‘situation modification’ (Duckworth et al. 2014, 2016). Note that situation modification in Gross’ (2015) original emotion regulation model is taken to mean ‘taking actions that directly alter a situation in order to change its emotional impact’ (p. 9), for example through placing an object that elicits a negative emotion from view. This definition does not seem to fully apply to the behaviours that we observed in the young children in our study. Yet, in the specific context of self-control, Duckworth et al. (2016) define situation modification more broadly, encompassing any changes in a person’s physical position in relation to the object that is to be avoided to facilitate self-control. Examples given include placing an alarm clock at the other end of the room so that a more effortful action is required to switch it off (getting out of the bed) than when it is placed on the nightstand. These examples are more closely aligned to what we have taken to mean ‘situation modification’, given that the young children we observed placed their hands under the table or behind their backs – thus changing the action required to reach to the object through changing their immediate physical situation.

The current results suggest that visual attention distraction and hands withholding behaviours may be strategic at this age, if we define strategic behaviour as children’s intentional actions aimed at succeeding on a task. A similar meaning was ascribed to the term ‘strategic behaviour’ in a recent micro-genetic study examining the way in which 13- and 18-month-old infants chose to descend a novel staircase (Berger et al. 2015). Following from this definition, we do not consider strategy use to be a fully conscious or metacognitive process at this age, but rather understand the term to mean the way young children approach a task or challenge with the intention of succeeding on that task or mastering that challenge.

At least one alternative to the idea that the behaviours that we observed are strategic in nature at this age needs to be considered. Children who manage to delay may simply have had more time to show any type of behaviour than children who fail, and as a consequence, the probability of occurrence may go up for each of the behaviours. For example, children who wait longer may become bored and thus start exploring the room. Previous experimental results by Mischel and colleagues (Mischel et al. 1989) counter this position, as do our own data on time effects. Indeed, if these behaviours occurred as a random by-product of having more time, one may expect them to emerge relatively late in the delay process. We observed the opposite: these behaviours were initiated very rapidly by successful children, within the first four to 8 seconds on average. In fact, about half to two thirds of successful children showed hands withholding behaviour at  $t = 1$  on both tasks, and about one quarter to one third showed visual attention distraction at  $t = 1$ . We incidentally noted that children would move their hands under the table during the instruction of the task, when the reward was not yet present on the table. These children thus understood the task already during the course of the instruction and were able to apply effective strategies immediately. Notwithstanding these findings, the current data are not suited for determining cause-effect relations. Experimental work is required in which, for example, the occurrence of hands withholding behaviours is manipulated through instruction.

The rapid initiation of the self-distraction and hands withholding strategies that we observed conflicts with previous studies showing much longer latencies in young children (Cole et al. 2011; Manfra et al. 2014). The discrepancy between our data and those of Cole and colleagues is likely due to a difference in methodology: whereas we coded when children looked away for

the first time by coding a single look away as self-distraction to obtain our latency measure, Cole and colleagues coded whether children became fully absorbed in self-distraction in 15-s epochs. Thus, in the latter study, self-distraction was full engagement with another toy, rather than simply looking away for a brief period of time. The coding approach of Manfra and colleagues seems closer to our own, and the discrepancy in results may be due to either a mean difference in age between studies (mean age four versus 3 years, respectively) or a difference in set-up. In the study by Manfra and colleagues, a large toy was placed on the ground, whereas a small reward was placed on the table in the current study. In our study, children thus had the table and their seat available as tools to help provide a physical ‘barrier’ between the object and themselves; this set-up may thus have provided them with more affordances for rapidly engaging in hands withholding behaviours than the set-up of the study by Manfra and colleagues. Although clearly requiring further testing, these findings tentatively speak to the situational nature of very early self-control.

Moreover, the results of the time effects analyses as plotted in Figs. 3a and B, show that, after an initial short phase characterized by primarily (visual) focus on the reward, visual attention distraction occurred increasingly more often, both alone and in combination with hands withholding. After 10–30 s, these two behaviours remained the two most probable behaviours to occur until the delay time was over, at 60 s, in both tasks. These findings tie in with recent studies showing that whether young children manage to wait for the first 10–20 s of the delay time or not is particularly predictive of later developmental outcomes, which suggests that something key happens in the first phase of the delay time (Friedman et al. 2011; Watts et al. 2018). Our findings suggest that what happens early on is the recruitment of effective visual attention distraction and hands withholding strategies. A logical hypothesis based on these findings, which could be tested in future studies, is that whether or not children manage to detect conflict and recruit such behavioural strategies very early on during the delay time is the core individual differences factor which is predictive of later outcomes.

Findings regarding the association between teacher reports of children’s self-control and the observed behaviours suggest that the latter have relevance beyond the specific delay of gratification test situation, although results between the snack and gift delay task were somewhat contradictory and effects did not hold for parent-rated self-control. First, children rated higher on self-control by their teachers showed a more rapid shift from combining visual attention distraction with hands withholding to visual attention distraction only over time on the snack delay task, compared to children lower on teacher-rated self-control. Although it is difficult to pin-point the meaning of this finding precisely, it may suggest that although many children rely strongly on the combined – and most controlled – strategy at the outset of the delay time (see Fig. 4), children with relatively high teacher ratings of self-control do not need to maintain this strategy over time to succeed on the task. Second, children rated higher on self-control by their teachers showed more hands withholding behaviours throughout than children lower on teacher-rated self-control on the gift delay task, and there were no significant time by teacher-rated self-control interactions. Thus, children with high teacher-rated self-control were more able to use a successful situational strategy early and to maintain that strategy over time, compared to children with lower teacher-rated self-control. The adult literature shows that adults high on trait self-control tend to identify a self-control conflict early (Gillebaarts et al. 2015), and pro-actively avoid situations that challenge self-control altogether (Ent et al. 2015). To the best of our knowledge, this is the first study showing

that individual differences in self-control in very young children relate to both the timing and occurrence of specific (situational) strategies.

It is of note that findings only occurred in relation to teacher- and not parent-rated behaviour. Previous studies have shown that young children's temperament and behavioural reports are only weakly to moderately interrelated between teachers and parents (Heyman et al. 2018; Phillips and Lonigan 2015; correlation in the current study:  $r = .21$ ,  $p = .066$ ). Both differences in informant perception of child behaviour and differences in actual behaviours that children display across contexts may explain the lack of a strong association. We speculate that the context in which the tasks were administered may in particular explain the discrepancy in findings related to parent and teacher reports in the current study. In our study, 87% of children were assessed in their day care center, while only 13% were assessed at home. Potentially, teacher reports on child behaviour were more closely aligned to the child's actual behaviour during the assessment than parent reports, as the former were observed in the same setting for the majority of cases.

### Future directions

An important question for future research is why some children demonstrate the visual attention distraction and hands withholding behaviours described in this study at such a young age, while others are not. What is the role of the social context in the emergence of these early individual differences? Lecuyer and Houck (2006) found that mothers' active attempts to direct their one-year-olds' attention away from a prohibited object were predictive of a longer delay of gratification time at the age of five. Similar findings were obtained in a cross-sectional study on mother-child interaction during a delay of gratification situation in 30-month-olds (Putnam et al. 2002). The next question is whether, in an experimental study, caregivers can be trained to effectively model situational and attentional strategies for self-control for their young children. If so, this would open up new avenues to provide parents and other caregivers with concrete advice on ways to promote self-control.

Moreover, although we investigated general time effects on occurrence of behaviours shown during delay, the current study did not address more specific dynamics over time. In particular, it would be interesting to investigate the transitional probabilities from one behavioural constellation to another and the individual differences factors that predict these probabilities. These types of analyses, such as Hidden Markov modelling, are highly promising techniques for further unravelling the dynamics of early childhood self-regulation processes (see for example, Lunkenheimer et al. 2017). The challenge with these analyses though is that they require a relatively large number of time points per child as well as a large sample of children. For example, Lunkenheimer and colleagues successfully modelled transitional probabilities using Hidden Markov modelling with second-by-second codes of 4–6-min parent-child interactions.

Finally, the current study findings and those of Friedman et al. (2011) and Watts et al. (2018) suggest that key individual differences in self-control are evident very early on in the delay time in young children. Future studies are required to unravel the order of occurrence of behaviours in the first part of the delay time, as well as the role of the strength of the emotional response (for example, through incorporating physiological measures), instruction (length, verbal complexity) and affordances of the test situation (such as the physical lay-out of the test situation) on occurrence on these very early emerging delay behaviours.

## Limitations

A number of limitations need to be considered when interpreting the findings from the current study. First, the environment in which the tests were conducted was non-standardized, as tests were all administered in the natural environment of the child's home or day-care, rather than in a controlled lab setting. For example, while some children were seated on a high children's chair at a high table that they could not climb out of, others were seated at a low chair and table from which they could easily stand up. Moreover, some children had their hands under the table already by the time the test instruction began, while others did not. We cannot rule out that these circumstances may have impacted on the behaviours children showed during delay. Thus, the current findings need to be interpreted with caution and require replication in a more controlled setting, in which the child's position at the outset of the delay time is also fully standardized. Second, although the sample size of the present study can be regarded as rather substantial for an observational multilevel study (Hox et al. 2018), a larger sample of children would have been preferable. Across the analyses we conducted, sample size seemed to influence the statistical significance of the effects of time on behavioural occurrence, while results regarding the associations between the observed behaviours and caregiver ratings of self-control appeared relatively robust. Also, it was not possible to conduct analyses predicting task success by latency of each of the behaviours, as the behaviours of interest did not occur in the majority of children who failed the task (and hence latency was missing for them; see Table 2). Third, the current study did not allow teasing apart effects of the type of reward (gift vs. snack) and task order, as task order was not counterbalanced. Fourth, delay time was relatively short in this study at 60 s, and children who succeeded on each task may have failed with a longer delay. The task was relatively easy for the older children in the sample, and we used a non-standard very short version of the ECBQ to allow data pooling across two- and three-year-olds. Thus, further studies are needed in which reward type and administration order are counterbalanced, sample size and delay time are increased, and different rewards are chosen to enhance the level of challenge involved for children aged 3 years and older.

## Conclusion

To the best of our knowledge, the current study is one of the first studies to provide a detailed view on the time course of two- and three-year-olds behaviours during delay of gratification. The study reveals that some children this young are already able to use various strategies to control their intention to reach for the reward, and rely strongly on both visual attention distraction and hands withholding behaviours. The latter type of behaviour was frequently observed at the beginning of the delay time, suggesting that these young children initiated effective self-control strategies very rapidly. Findings from the current study need to be replicated with a larger sample in a more standardized setting. If replicated, the current work may provide a starting point for research into interventions and early childhood education targeted at stimulating delay of gratification ability. In addition, the current findings call for further developmental studies into the temporal dynamics of strategy use during delay of gratification in different (risk) groups and under various task conditions, to further pinpoint the individual and contextual factors that shape the fine-grained dynamics of strategy use for self-control early in life.

**Acknowledgements** We are very grateful to all children, parents, and teachers for taking part in the current study. We also thank Lea Aarts, Stefanie van der Burgt, Leonie Stoevenbelt, Veronique van Dam, Daphne Leemans, and Marloes van der Hagen for their help with data collection.

**Funding** This study was funded by the Dutch research council NWO (grant number grant number 411–20–442).

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## Appendix A: Coding scheme of delay of gratification strategies

Behavioral modality	Behaviour	Codes
Visual	Focus	Looks towards object
	Withhold	Eyes are closed
	Distract*	Looks left or right of the object
		Looks in the opposite direction of the object
		Looks upwards
		Looks at his/her own hands
		Looks at a cue in the environment
Looks at another person		
Looks downwards		
Verbal	Focus	Talks about the object to someone else
		Talks about what he/she does (action) regarding the object to someone else
		Asks questions about the object to someone else
		Talks about the object to him/ herself
		Talks about what he/she does (action) regarding the object to him/ herself
	Withhold	Asks questions about the object to him/ herself
		Repeats the rule regarding the task to someone else
		Repeats the rule regarding the task to him/herself
	Distract	Making contact with others (in general)
		Talks about object-unrelated actions to someone else
		Asks general questions to someone else
		Sings or hums
		Makes noises (no distinguishable language)
		Makes mouth movements (without sound)
		Talks in general to him/ herself
Talks about what he/she does in general to him/ herself		
Uses unintelligible speech, such as muttering		
Motor - Hands	Does not speak	
	Focus	Almost touching the object
		Reaching or pointing towards the object
		Hands are (relaxed) above the table
	Withhold*	Hands are placed on top of one another
		Holding own hands
		Hands are underneath the table
		Sits on own hands
	Distract*	Head rests on the hands
		Touches own face or mouth
<i>Failure*</i>	Fine motor movements (e.g., fiddles with clothes)	
	Touches/holds another object (not task- related)	
	Low: The child touches, holds, or shakes the object Medium: The child unpacks the object	

Behavioral modality	Behaviour	Codes
Motor - Head	Focus	High: The child uses the object Faces the object directly
	Distract	Faces the object closely, within a 45° angle Face is left of the object Face is right of the object Face is directed upwards Face is directed downwards Face is directed opposite of the object Head is resting on the table
Motor - whole body	Focus	Body is turned towards the object
	Distract*	Body is turned sideways (left or right of the object) Body is turned completely around Stands besides the table/chair Rocks or dances with its body The child is away from the table

\*Behavioural categories that were used in the analyses

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Berger, S. E., Chin, B., Basra, S., & Kim, H. (2015). Step by step: A microgenetic study of the development of strategy choice in infancy. *British Journal of Developmental Psychology*, *33*, 106–122. <https://doi.org/10.1111/bjdp.12076>.
- Bos, J., & Steenbeek, H. W. (2009). *Mediacoder 2009: A simple application for coding behaviour within media files*. Groningen: University of Groningen.
- Bridgett, D. J., Burt, N. M., Edwards, E. S., & Deater-Deckard, K. (2015). Intergenerational transmission of self-regulation: A multidisciplinary review and integrative conceptual framework. *Psychological Bulletin*, *141*, 602–654. <https://doi.org/10.1037/a0038662>.
- Carlson, S. M., & Beck, D. M. (2009). Symbols as tools in the development of executive function. In A. Winsler, C. Fernyhough, & I. Montero (Eds.), *Private speech, executive functioning, and the development of verbal self-regulation*. New York: Cambridge University Press.
- Cole, P. M., Tan, P. Z., Hall, S. E., Zhang, Y., Crnic, K. A., Blair, C. B., & Li, R. (2011). Developmental changes in anger expression and attention focus: Learning to wait. *Developmental Psychology*, *47*, 1078–1089. <https://doi.org/10.1037/a0023813>.
- Duckworth, A. L., Gendler, T. S., & Gross, J. J. (2014). Self-control in school-aged children. *Educational Psychologist*, *49*, 199–217. <https://doi.org/10.1080/00461520.2014.926225>.
- Duckworth, A. L., Gendler, T. S., & Gross, J. J. (2016). Situational strategies for self-control. *Perspectives on Psychological Science*, *11*, 35–55. <https://doi.org/10.1177/1745691615623247>.
- Efron, B., & Tibshirani, R. J. (1993). An introduction to the bootstrap. *Monographs on Statistics and Applied Probability*. Dordrecht, the Netherlands: Springer Science+Business Media.
- Ent, M., Baumeister, R. F., & Tice, D. M. (2015). Trait self-control and the avoidance of temptation. *Personality and Individual Differences*, *74*, 12–15. <https://doi.org/10.1016/j.paid.2014.09.031>.
- Friedman, N. P., Miyaki, A., Robinson, J. L., & Hewitt, J. K. (2011). Developmental trajectories in toddlers' self-restraint predict individual differences in executive functions 14 years later: A behavioural genetic analysis. *Developmental Psychology*, *47*, 1410–1430. <https://doi.org/10.1037/a0023750>.
- Fujita, K. (2011). On conceptualizing self-control as more than the effortful inhibition of impulses. *Personality and Social Psychology Review*, *15*, 352–366. <https://doi.org/10.1177/1088868311411165>.

- Gillebaarts, M., Schneider, I. K., & De Ridder, D. T. D. (2015). Effects of trait self-control on response conflict about healthy and unhealthy food. *Journal of Personality, 84*, 789–798. <https://doi.org/10.1111/jopy.12219>.
- Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology, 2*, 271–299. <https://doi.org/10.1037/1089-2680.2.3.271>.
- Gross, J. J. (2015). Emotion regulation: Current status and future prospects. *Psychological Inquiry, 26*, 1–26. <https://doi.org/10.1080/1047840X.2014.940781>.
- Hedeker, D. (2003). A mixed-effects multinomial logistic regression model. *Statistics in Medicine, 22*, 1433–1446. <https://doi.org/10.1002/sim.1522>.
- Heyman, M., Poulakos, A., Upshur, C., & Wenz-Gross, M. (2018). Discrepancies in parent and teacher ratings of low-income preschooler's social skills. *Early Childhood Education and Care, 188*, 759–773. <https://doi.org/10.1080/03004430.2016.1236257>.
- Hofmann, W., Friese, M., & Strack, F. (2009). Impulse and self-control from a dual-systems perspective. *Perspectives on Psychological Science, 4*, 162–176. <https://doi.org/10.1111/j.1745-6924.2009.01116.x>.
- Hofmann, W., Baumeister, R. F., Förster, G., & Vohs, K. D. (2012). Everyday temptations: An experience sampling study of desire, conflict, and self-control. *Journal of Personality and Social Psychology, 102*, 1318–1355. <https://doi.org/10.1037/a0026545>.
- Hox, J., Moerbeek, M., & Van de Schoot, R. (2018). *Multilevel analyses: Techniques and applications*. New York: Taylor & Francis Ltd..
- Kochanska, G., Murray, K. T., Jacques, T. Y., Koenig, A. L., & Vandegeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Development, 67*(2), 490–507. <https://doi.org/10.1111/j.1467-8624.1996.tb01747.x>.
- Kochanska, G., Murray, K., & Coy, K. C. (1997). Inhibitory control as a contributor to conscience in childhood: From toddler to early school age. *Child Development, 68*, 263–277.
- Kochanska, G., Murray, K. T., & Harlan, E. T. (2000). Effortful control in early childhood: Continuity and change, antecedents and implications for social development. *Developmental Psychology, 36*, 220–232. <https://doi.org/10.1037//0012-1649.36.2.220>.
- Kopp, C. B. (1982). Antecedents of self-regulation: A developmental perspective. *Developmental Psychology, 18*, 199–214. <https://doi.org/10.1037/0012-1649.18.2.199>.
- Kottner, J., & Dassen, T. (2008). Interpreting interrater reliability coefficients of the Braden scale: A discussion paper. *International Journal of Nursing Studies, 45*, 1238–1246. <https://doi.org/10.1016/j.ijnurstu.2007.08.001>.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159–174.
- Lecuyer, E., & Houck, G. M. (2006). Maternal limit-setting in toddlerhood: Socialization strategies for the development of self-regulation. *Infant Mental Health Journal, 27*(4), 344–370. <https://doi.org/10.1002/imhj.20096>.
- Lunkenheimer, E., Ram, N., Skowron, E. A., & Yin, P. (2017). Harsh parenting, child behavior problems, and the dynamic coupling of parents' and children's positive behaviors. *Journal of Family Psychology, 31*, 689–698. <https://doi.org/10.1037/fam0000310>.
- Manfra, L., Davis, K. D., Ducenne, L., & Winsler, A. (2014). Preschoolers' motor and verbal self-control strategies during a resistance-to-temptation task. *The Journal of Genetic Psychology, 175*(4), 332–345. <https://doi.org/10.1080/00221325.2014.917067>.
- Metcalfe, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: dynamics of willpower. *Psychological Review, 106*, 3–19.
- Mischel, W., & Baker, N. (1975). Cognitive appraisals and transformations in delay behavior. *Journal of Personality and Social Psychology, 31*, 254–261.
- Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science, 244*, 933–938 Retrieved from: <http://www.jstor.org/stable/1704494>.
- Moffitt T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., Houts, R., ... Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*. [www.pnas.org/cgi/doi/10.1073/pnas.1010076108](http://www.pnas.org/cgi/doi/10.1073/pnas.1010076108)
- Mulder, H., Hoofs, H., Verhagen, J., Van der Veen, I., & Leseman, P. P. M. (2014). Psychometric properties and convergent and predictive validity of an executive function test battery for two-year-olds. *Frontiers in Psychology, 5*:733. <https://doi.org/10.3389/fpsyg.2014.00733>.
- Nigg, J. T. (2017). Annual research review: On the relations among self-regulation, self-control, executive functioning, effortful control, cognitive control, impulsivity, risk-taking, and inhibitory for developmental psychopathology. *Journal of Child Psychology and Psychiatry, 58*, 361–383. <https://doi.org/10.1111/jcpp.12675>.
- Peake, P. K., Mischel, W., & Hebl, M. (2002). Strategic attention deployment for delay of gratification in working and waiting situations. *Developmental Psychology, 38*, 313–326.

- Phillips, B. M., & Lonigan, C. J. (2015). Child and informant influences on behavioral ratings of preschool children. *Psychology in the Schools, 47*, 374–390. <https://doi.org/10.1002/pits.20476>.
- Putnam, S. P., Garstein, M. A., & Rothbart, M. K. (2006). Measurement of fine-grained aspects of toddler temperament: The Early Childhood Behavior Questionnaire. *Infant Behavior & Development, 29*, 386–401.
- Putnam, S. P., Spritz, B. L., & Stifter, C. A. (2002). Mother-child coregulation during delay of gratification at 30 months. *Infancy, 3*, 209–225. doi:10.1207/S15327078IN0302\_6
- Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at 3–7 years: The Children's Behavior Questionnaire. *Child Development, 72*, 1394–1408.
- Schlam, T. R., Wilson, N. L., Shoda, Y., Mischel, W., & Ayduk, O. (2013). Preschoolers' delay of gratification predicts their body mass 30 years later. *Journal of Pediatrics, 162*, 90–93.
- Sethi, A., Mischel, W., Aber, J. L., Shoda, Y., & Rodriguez, M. L. (2000). The role of strategic attention deployment in development of self-regulation: Predicting preschoolers' delay of gratification from mother-toddler interactions. *Developmental Psychology, 36*, 767–777. <https://doi.org/10.1037/0012-1649.36.6.767>.
- Shoda, Y., Mischel, W., & Peake, P. K. (1990). Predicting adolescent cognitive and self-regulatory competencies from preschool delay of gratification: Identifying diagnostic conditions. *Developmental Psychology, 26*, 978–986.
- Vaughn, B. E., Kopp, C. B., Krakow, J. B., Johnson, K., & Schwartz, S. S. (1986). Process analyses of the behavior of very young children in delay tasks. *Developmental Psychology, 22*, 752–759.
- Watts, T. W., Duncan, G. J., & Quan, H. (2018). Revisiting the marshmallow test: A conceptual replication investigating links between early delay of gratification and later outcomes. *Psychological Science, 29*, 1159–1177. <https://doi.org/10.1177/0956797618761661>.
- Wilson, A. C., Lengua, L. J., Tininenko, J., Taylor, A., & Trancik, A. (2009). Physiological profiles during delay of gratification: Associations with emotionality, self-regulation, and adjustment problems. *Journal of Applied Developmental Psychology, 30*(6), 780–790. <https://doi.org/10.1016/j.appdev.2009.05.002>.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Affiliations

Hanna Mulder<sup>1</sup> · Heleen van Ravenswaaij<sup>2</sup> · Josje Verhagen<sup>1</sup> · Mirjam Moerbeek<sup>3</sup> · Paul P. M. Leseman<sup>1</sup>

<sup>1</sup> Department of Special Education, Cognitive and Motor Disabilities, Utrecht University, Heidelberglaan 1, 3584 CS Utrecht, the Netherlands

<sup>2</sup> Biomedical Sciences, University Medical Centre, Utrecht, the Netherlands

<sup>3</sup> Department of Methodology and Statistics, Utrecht University, Utrecht, the Netherlands