Examining delay of gratification in healthy aging

Bidhan Lamichhane1*, Elisa Di Rosa1*, Leonard Green1, Joel Myerson1 and Todd S. Braver1

1 Department of Psychological and Brain Sciences, Washington University in St. Louis (US)

*These authors contributed equally to the work

Abstract

Delay of gratification (DofG) refers to the capacity to forego an immediate reward in order to receive a more desirable reward later. As a core executive function, it might be expected that DofG would follow the standard pattern of age-related decline observed in older adults for other executive tasks. However, there actually have been few studies of aging and DofG, and even these have shown mixed results, suggesting the need for further investigation and new approaches. The present study tested a novel reward-based decision-making paradigm enabling examination of age-related DofG effects in adult humans. Results showed that older adults earned fewer overall rewards than young adults, both before and after instruction regarding the optimal DofG strategy. Prior to instruction, learning this strategy was challenging for all participants, regardless of age. The finding of age-related impairments even after strategy instruction indicated that these impairments were not due to a failure to understand the task or follow the optimal strategy, but instead were related to self-reported difficulty in waiting for delayed rewards. These results suggest the presence of age-related changes in DofG capacity and highlight the advantages of this new experimental paradigm for use in future investigations, including both behavioral and neuroimaging studies.
1. Introduction

The aging of the worldwide population has focused attention on the question of how economic decision-making capacity changes across the adult lifespan (for reviews, see Samanez-Larkin and Knutson, 2015; Sparrow and Spaniol, 2016; Spreng et al., 2016). Intertemporal choice, a key aspect of economic decision-making, involves choices whose outcomes may occur at different points in time. Intertemporal choice in older adults is of interest not only because of its obvious financial consequences (e.g., investment decisions), but also because it is widely thought to index the capacity for self-control. More specifically, delay of gratification (DofG) refers to the capacity to forego an immediate reward in order to receive a more desirable reward later (Mischel, 1974; Mischel al., 1989) and is typically treated in the developmental and neuropsychological literatures as a core “executive function” (Mischel et al., 2010; Casey et al., 2011; Zelazo and Carlson, 2012).

Typically, DofG tasks are lumped together with other intertemporal choice tasks, such as temporal discounting tasks, which have been examined frequently in the human neuroscience literature, and which appear to rely on neural circuitry whose functions (e.g., the interactions between the prefrontal cortex and subcortical dopaminergic pathways) decline with age (Wittmann et al., 2007; Achterberg et al. 2016; Schüller et al. 2019). However, the procedure used in DofG tasks is different from that used in temporal discounting tasks (Reynolds and Schiffbauer 2005; Göllner et al., 2018) in that DofG requires repeated decisions regarding whether to continue waiting for a larger delayed reward or instead to obtain a currently available but smaller reward. Notably, although temporal discounting tasks have been frequently studied in both younger and older adults (for recent evidence, see Sasse et al., 2017; Han et al., 2018), there has been surprisingly little study of age-related differences in DofG, making it a ripe target for further investigation.

Because DofG is considered a high-level cognitive function, relying on brain structures that have expanded most recently in evolution (i.e., prefrontal cortex), it might be reasonable to expect that it would have developed late during phylogeny, yet studies conducted in many animal species have demonstrated that DofG capacity is not unique to humans. Indeed, several lines of evidence have shown that a number of animal species (e.g., chimpanzees, dogs, parrots) exhibit the capacity to endure long periods of delay in return for a larger reward payoff (e.g., Koepke et al., 2015; Brucks et al., 2017; Beran and Hopkins, 2018; for a recent review, see Flessert and
Beran, 2018). In terms of ontogenetic development, it might be predicted that DofG capacity would show the inverse U-shaped pattern characteristic of other high-level cognitive functions (e.g., executive function; Zelazo et al., 2004), and therefore that DofG capacity, which is immature in children, would decline in older adulthood.

The question as to whether adults’ DofG capacity declines in older age is still open because most of the research has been conducted using children, and with paradigms (e.g., the well-known “Marshmallow Task”; Mischel, 1974; Mischel et al., 1988) that are not conducive to research with adults and/or would not be suitable for use in neuroimaging studies. In fact, to the best of our knowledge, although some research has been conducted employing self-report questionnaires (for a review see Drobetz et al., 2012), only a very few studies have examined DofG capacity in older adults using experimental tasks. Specifically, Forstmeier et al. (2011) designed and validated a DofG task for adults (the DofG-A) that included different types of rewards (snacks, real money, hypothetical money, and magazines) that were chosen so as to be meaningful for older adults. However, their study focused only on individuals aged 60 years and older, and as such did not provide any information regarding potential differences in DofG capacity between young and older adults. This same task was employed more recently by Göllner et al. (2018), who tested participants between the ages of 9 and 101 years, divided into five age groups (9–14, 18–25, 35–55, 65–80, 80+). They also employed a delay discounting task and reported significant age differences in that capacity. Age differences in DofG, however, could be found only with an alpha level of .10. Thus, the question of whether DofG capacity changes during aging is still open and in need of further investigation.

Accordingly, the goal of the present study was to develop an experimental paradigm to assess DofG capacity in younger and older adults. In order to maximize face validity and ensure continuity with prior literature, we adapted an existing paradigm (Hackenberg and Axtell, 1993) that showed sensitivity to DofG-relevant experimental variables in both adult humans and in nonhuman species.

2. Material and Methods
2.1 Participants

Forty-one healthy participants were recruited: 21 younger adults (age 18-28 years; mean=19.9, SD=2.3; 6 Males) and 20 older adults (age 65-83 years, mean=73.1, SD=6; 6 Males), with the following self-reported demographic breakdown: White/Caucasian (9 younger, 14 older), African-American (5 younger, 4 older), Asian (6 older), Hispanic (1 older), and multiracial (1 older). Exclusion criteria were the presence or history of neurological or psychiatric disease, and the use of neurological or psychiatric medications. All participants had normal or corrected-to-normal vision and provided written informed consent. Participants were reimbursed $10/hour for their participation, plus any monetary earnings during the task.

2.2 Experimental task and procedure

Participants performed an economic decision-making task adapted from Hackenberg and Axtell (1993) in which the amount of monetary reward they earned depended on their choices. On each trial, participants had to select one of two response options indicated by blue and yellow boxes presented at the left and the right side of the monitor, respectively (see Figure 1) and press the corresponding button on a response box. Each of the two options offered the possibility of obtaining a reward ($0.10), but they paid off according to different schedules. Choice of the option associated with the yellow box resulted in subsequent rewards being available according to a progressive-interval (PI) schedule, starting from 2 s after the beginning of the trial and increasing with fixed increments of 4 s following each subsequent PI reward (i.e., across trials). In contrast, choice of the blue option resulted in rewards being available according to a fixed-interval (FI) schedule (i.e., every 40 s after the beginning of the trial). Critically, choice of the blue (FI) option also reset the value of the progressive-interval schedule for the yellow option back to its initial (2 s) delay. Prior to performing the task, participants were informed regarding the different schedules associated with the two options. Participants also were informed that when options became available, they were free to take their time to make their choice. That is, the PI or FI timer would stop, and the options would remain on the screen until the participant made their choice. The purpose of this instruction was to encourage participants to be sure of their decisions and also to minimize the effects of age-related differences in choice latency.
To illustrate the procedure, Figure 1 depicts two hypothetical sequences of events: Sequence A (shown on the left) depicts what happens when a participant selects the FI schedule option on one trial and the PI option on the next trial; Sequence B (shown on the right) depicts what happens when a participant selects the PI schedule option on two successive trials. The beginning of each trial occurs with the blue box and the yellow box indicating a response is required (“Press to start trial”); once either response button is pressed, the timers for the PI and FI schedules are started. In Sequence A, the participant chose to forego the available reward on the PI schedule (yellow box). Thus, the critical events occur when the PI schedule delay has ended and its reward is available but the participant chose to keep waiting for the FI reward (i.e., a DofG trial) by pressing the appropriate response button for the FI schedule (blue box; Sequence A, Trial N, “Choose wait”). At the end of the FI schedule delay (40 s), the FI reward becomes available and is obtained by pressing the appropriate response button (blue box). When the response for the FI reward is selected, the reward feedback message is displayed, the trial ends, and critically, the PI schedule delay is reset to its initial value, such that the PI reward is now available after 2 s (Sequence A, Trial N+1). In Sequence B, an alternative trial sequence is shown in which the participant chooses the PI reward. Here, when the PI schedule delay has ended on the trial (e.g., after, say, 18 s), the participant selected the available reward on the PI schedule (yellow box) by pressing the corresponding response button. In this case, the reward feedback message is displayed, and the trial ends, but now the PI schedule delay is increased by 4 seconds on the next trial (e.g., to 22 s; Sequence B, Trial N+1).

It is to be noted that on all trials, prior to the end of the PI schedule delay, both boxes present the same message (“press to keep waiting”) with a 50% probability every 2 s (i.e., forced wait events, see Figure 1). Likewise, on trials in which the PI schedule delay ends but participants choose to forego the PI reward (i.e., DofG trials), the option reoccurs every 2 s with 50% probability, requiring participants to make a choice between choosing to keep waiting (i.e., choose wait events, engaging DofG) or instead allowing them the opportunity to “defect” and select the available reward. These event were included to match the response and attentional demands present on PI and FI (DofG) trials, with both event types (forced wait, choose wait) occurring probabilistically, to reduce the predictability of responding and maintain participants’ attention throughout the trial.
Participants performed four task Blocks, each approximately 12 minutes in duration. The first two Blocks were defined as “learning Blocks” because participants were asked to develop their own preferred strategy in order to gain as much money as they could. At the end of each Block, participants were informed regarding the amount of reward they had earned on that Block and that this amount was being added to their total compensation. In the third Block, participants received additional instructions. Specifically, they were asked to follow a precise strategy of choosing the reward associated with the yellow box (i.e. the one following the PI schedule) for four trials in a row and then waiting for the reward associated with the blue box (i.e., the one following the FI schedule), choosing it, repeating this same sequence of choices (4 yellow followed by 1 blue) for the entire Block.

This trial sequence actually was the optimal strategy because if followed exactly, it would result in earning the maximum possible amount of reward available for the Block ($5.10). However, participants were not informed that this strategy was optimal. In the fourth and final Block, the participants were told to follow whatever strategy they wished, including, if they preferred, continuing with the one they were instructed to use in the third Block, and were told, “some individuals feel that the strategy followed in the third Block was the best one.”
Figure 1. The experimental DofG paradigm and trial sequences. Left side (Sequence A): sequence when choosing the FI reward (DofG option) in the first of two successive trials (N); the PI schedule delay interval resets to its minimum value (2 sec) on the second trial (N+1). The repeated requirement to forego an available PI reward (“choose wait”; engage DofG) is highlighted. Right side (Sequence B): sequence when choosing the PI reward (first available, impulsive option) on two successive trials (N, N+1); reward delay interval increases.

2.3 Post-task questionnaire

After completing the four Blocks, participants were asked to fill out a post-task strategy questionnaire. The questionnaire asked the participants to describe their strategy or pattern of responding for each Block and to provide Likert-scale ratings (1 to 5, where 1 is “not at all” and 5 is “a lot”) regarding the following: a) How much influence did you feel like your strategy had on
the total amount of money you earned? b) How difficult was it when you had to wait for a reward when none were available? and c) How difficult was it to wait for the yellow reward when the blue reward was already available? For the first, second, and fourth Blocks, participants also had to provide Likert-scale ratings for the question: How much did you vary your strategy to try to earn more money? For the third Block, they were asked to: a) indicate if you tried to maintain the suggested strategy during the entire Block (yes or no); and b) rate how closely you were able to stick to that strategy. Finally, they were asked to describe what they thought the ideal strategy was and why they thought it worked the best.

2.4 Data analysis

Task performance was assessed by calculating the percentage of the maximum possible reward that a participant obtained in each Block. The normality of the distribution for each block and for the Likert-scale ratings on the post-task questionnaire was assessed using the Kolmogorov-Smirnoff test, and results indicated that only performance in Block 2 was normally distributed ($p = .09$) whereas the other measures were not (all $ps < .05$). Levene's test was used to assess the equality of variances, and significant differences were observed for Blocks 3 and 4 ($p < .001$). Therefore, within-group differences were investigated using the Wilcoxon signed-rank test, and between-groups differences were investigated using the Mann–Whitney U test. Based on a Bonferroni correction for multiple comparisons, the significance level was set equal to .003.

Analyses of the post-task questionnaire were primarily exploratory and conducted to guide hypothesis generation for follow-up studies. Correlations between questionnaire responses and task performance were assessed with Spearman's rank correlation coefficient. We report results that meet conventional statistical significance levels ($p < .05$), but we did not apply multiple comparisons correction; thus, statistical inference should proceed cautiously.

3. Results

Performance of both groups in each of the four task Blocks is reported in Table 1. Results of the within-group statistical analyses indicated that younger adults’ performance significantly improved in Block 3 compared with both Blocks 1 and 2 ($Z = - 4.02$, $p < .0001$, and $Z = - 3.92$, $p < .0001$, respectively).
This performance improvement was maintained in Block 4, where performance again differed significantly from both Block 1 \((Z = -4.02, p < .0001)\) and Block 2 \((Z = -3.93, p < .0001)\). Young adults’ performance in Block 4 did not differ from that in Block 3 \((Z = 0.82, p = .408)\).

### Table 1: Task performance (% of maximum possible reward obtained) of young and older adults in each of the four task Blocks (Standard Deviations are given in parentheses).

<table>
<thead>
<tr>
<th>Group</th>
<th>Block 1</th>
<th>Block 2</th>
<th>Block 3</th>
<th>Block 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young</td>
<td>82.5</td>
<td>83.6</td>
<td>97.2</td>
<td>96.4</td>
</tr>
<tr>
<td>(N=21)</td>
<td>(12.8)</td>
<td>(14.3)</td>
<td>(2.2)</td>
<td>(4.0)</td>
</tr>
<tr>
<td>Older</td>
<td>72.7</td>
<td>70.5</td>
<td>93.4</td>
<td>82.7</td>
</tr>
<tr>
<td>(N=20)</td>
<td>(12.0)</td>
<td>(10.9)</td>
<td>(11.0)</td>
<td>(15.5)</td>
</tr>
</tbody>
</table>

A similar pattern of performance was observed in the older adults, with a significant improvement in Block 3 over performance in Blocks 1 and 2 \((Z = -3.81, p < .0001, \text{and } Z = -3.70, p < .0001, \text{respectively})\), and improved performance in Block 4 relative to Blocks 1 and 2 again was observed \((Z = -3.34, p < .002, \text{and } Z = -3.16, p < .003, \text{respectively})\). Although older adults’ performance, on average, decreased from 93% in Block 3 to 83% in Block 4, this change did not reach the Bonferonni-corrected significance threshold \((Z = 2.56, p = .01)\).

Between-group analyses indicated that although the age difference was not reliable in Block 1 \((U = -2.5, p = .012)\), a significant age difference was observed in Block 2 \((U = -3.39; p < .002)\). Notably, this age difference completely disappeared in Block 3 \((U = -0.94; p = .34)\), but then reappeared in Block 4 \((U = -3.16, p < .003)\).
Examining delay of gratification in healthy aging

Figure 2. Performance of both age groups across task blocks. Performance shown as the percentage of total maximum possible reward obtained in each block. YA = young adults (grey bars), OA = older adults (unfilled bars). Small circles indicate individual participants. Asterisks (+brackets) at the top of the figure indicate the significant block-related differences (Block 4 > Block 1; Block 4 > Block 2; Block 3 > Block 1; Block 3 > Block 2) present in both young and older adults. Asterisks (+arrows) at the bottom right of the figure indicate between-group differences (YA > OA in Blocks 2 and 4).

Analyses of the self-report questionnaires provided information regarding the source of potential decision-making challenges associated with the DofG task. Task performance in Block 3 was negatively correlated (Rho = -.32) with self-reported difficulty in DofG (i.e., the difficulty in waiting for the blue reward when the yellow reward was already available). In Block 4, performance was negatively related to a more general difficulty in waiting for both kinds of rewards (Rho = -.40) as well as to the degree of strategy variation employed (Rho = -.46). The pattern of correlations also suggests a positive association of level of performance with the perceived influence of the strategy being used: For Blocks 2, 3, and 4, Rho equaled .32, .36, and .63, respectively. Finally, in Block 4 participants who reported varying their strategy more also reported both more difficulty in waiting for rewards (Rho = .64) and more difficulty in making DofG responses (Rho = .52).
4. Discussion and Conclusion

The goal of the present study was to validate the utility and sensitivity of an experimental paradigm for detecting potential differences between younger and older adults in DofG capacity. A key feature of this paradigm is that in the first two Blocks, participants are free to explore and test different strategies to learn which earns more money, whereas in the third Block participants are tested to see whether they can follow the optimal strategy when instructed, even though they were not explicitly told that it is optimal. Finally, in the fourth Block, participants are again allowed to freely select their preferred decision-making strategy in order to assess the impact of prior learning on their performance, as well as to obtain a measure of asymptotic choice behavior and DofG capacity.

Several key findings emerged. First, learning the optimal DofG strategy from experience was challenging for all participants, regardless of age, as evidenced by the fact that performance in the first two Blocks was well below ceiling, but improved significantly in both groups following instruction. Second, significant age differences were observed in strategy learning and choice behavior. Specifically, in Block 2, older adults exhibited poorer performance than young adults, but this age difference disappeared in Block 3 after participants were given explicit instruction in the optimal DofG strategy, indicating that the age-differences do not reflect an inability to follow the complex optimal decision-making strategy that the task required (i.e., choose the PI option for 4 consecutive trials, then choose the FI option, and then repeat this whole sequence). Strikingly, however, an age difference reappeared in Block 4, even after practice following the instructed strategy in Block 3, suggesting that the observed age difference reflects either a memory deficit or a true decision-making strategy preference (i.e., a potential motivational shift) and/or an impairment in DofG capacity (i.e., a potential decline in self-control).

Findings from the post-task self-report measures suggest possible hypotheses regarding the source of individual and age-related variation in DofG capacity. Correlational analyses revealed possible associations of task performance with self-reported difficulty in waiting for rewards, and with waiting for the delayed reward associated with the FI option (i.e., with making DofG choices) in particular. In addition, participants who reported varying their strategy more tended to report greater difficulty in following a decision-making strategy requiring DofG capacity, and participants who felt that there was a clear effect of decision-making strategy on reward earning.
tended to earn more money. Together, these findings suggest that implementation of a strategy dependent on DofG capacity may indeed involve a difficult self-control challenge, and that this difficulty may prompt some individuals to explore alternative but suboptimal decision-making strategies.

The present results provide important initial validation of a new decision-making paradigm able to probe DofG capacity in both younger and older adults, although further study clearly is needed. First, it will be important to assess the relationship between DofG capacity and performance on other executive function tasks as well as other affective/motivational tasks that may tap into self-control. The recruitment of a broader age-span of participants (including a middle-aged group) would also be useful to directly demonstrate an inverted U-shaped function in DofG capacity. Secondly, comparing performance on our DofG task with performance on related tasks such as temporal discounting would be highly valuable and could reveal the extent to which such tasks involve common processes. It also will be important to determine whether the observed age differences in DofG performance reflect covarying neuropsychological changes or other confounding factors (e.g., health, income, education, IQ). Lastly, it is to be noted that in the present study we cannot disentangle the effect of instructions from the effect of practice/learning on choice performance. Recruiting a control group that performs the four task blocks without any DofG instruction would therefore be recommended in future studies.

Currently, our research group is conducting a follow-up neuroimaging study with this paradigm to determine its utility for revealing the neural basis of DofG capacity. The paradigm was explicitly adapted to be amenable for use in such contexts, with a trial-based design that enables clear contrasts between DofG and non-DofG events (i.e., between having to wait for one reward when another would be available sooner versus merely having to wait for a reward), while also allowing for examination of pre-decision neural activity dynamics. The use of this DofG task in a neuroimaging context could shed light on the neural mechanisms that give rise to successful DofG choices, as well as on distinctions between DofG and other components of reward-based and intertemporal decision-making, thereby advancing our understanding of potential sources of age-related change in decision-making behavior.
Acknowledgements

This work was supported by the National Institutes of Aging [NIA R21AG05820, R01AG058885, and R01AG043461]. EDR is funded by the Marie-Skłodowska-Curie Global Fellowship MOTIVAGEING, GA# 749084 (Horizon 2020). The authors are grateful to Sarah Adams and Jeremy Delgadillo for their assistance in data collection and processing.
Examining delay of gratification in healthy aging

References


