Motiv Emot (2011) 35:118–134
DOI 10.1007/s11031-011-9213-1

ORIGINAL PAPER

The delay of gratification test for adults: Validating a behavioral measure of self-motivation in a sample of older people

Simon Forstmeier · Reinhard Drobetz · Andreas Maercker

Published online: 8 April 2011

Abstract Most previous delay of gratification tests were developed for children and are inappropriate for application in adults. The authors therefore developed the Delay of Gratification Test for Adults (DoG-A), which includes four types of reward that are meaningful to adults, namely snacks, real money, hypothetical money, and magazines. Four subscores and two composite scores can be calculated. This study is the first to evaluate the DoG-A and to investigate its association with external variables. A community sample of 147 cognitively healthy participants aged between 60 and 94 years completed a questionnaire and cognitive tests measuring delay discounting, self-regulation, motivational self-concept, personality, wellbeing, and cognitive function. The intercorrelations of the subscales were low to medium and the internal consistency of the composite scores was moderate ($\alpha = .4$), indicating relative domain independence of the four reward types. The nomological net established by investigating the relations of the DoG-A with other constructs proved to be fairly meaningful. The correlations of all subscales with the delay discounting rate were significant and moderate. The Snacks subscale showed the most consistent pattern of results in terms of moderate positive correlations with self-reported motivation regulation, optimism, dutifulness, and deliberation. The Snacks subscale also correlated with various measures of wellbeing. A regression analysis showed that DoG Snacks remained a significant predictor of wellbeing when self-reported self-regulation and other variables were controlled. These findings indicate that the DoG-A yields an interpretable behavioral measure of self-motivation and offers a developmentally adequate extension of the delay of gratification paradigm for use with adults.

Keywords Delay of gratification · Self-motivation · Self-regulation · Self-control · Delay discounting · Older age · Wellbeing

S. Forstmeier (✉) · R. Drobetz · A. Maercker
Department of Psychology, Division of Psychopathology and Clinical Intervention, University of Zurich, Binzmuehlestr. 14/17, 8050 Zurich, Switzerland
e-mail: s.forstmeier@psychologie.uzh.ch

Introduction

Having to choose between an immediate reward and a greater future reward that may require the investment of time and effort is a recurrent challenge in life. Examples include saving money instead of caving into impulse buys, doing a disagreeable job for the gratification of a salary at the end of the month, resisting the temptation of sweets when dieting, stopping smoking to reduce the future health risks, and staying into study for university exams. This voluntary postponement of an immediate reward for a later but larger one has been termed *delay of gratification* (Mischel et al. 1989). The construct of delay of gratification (DoG), often interpreted as self-control or self-regulation, has attracted research interest for almost 60 years (Mischel and Ebbesen 1970). It was originally investigated in children between four and 6 years of age and, to this day, little research has been conducted with adults—almost none with older adults. The reason for this neglect may be rooted in the classic DoG paradigm of the "marshmallow test." Although this approach is appropriate for children, marshmallows are not appropriate rewards for adults. Yet adults face situations requiring them to choose between immediate and delayed gratification on a daily basis. Thus, there is a clear need for adequate methods of measuring DoG in adulthood and in old age.

Assessment of delay of gratification

In the classic DoG paradigm, the child is presented with a reward, such as a marshmallow or a cookie (Mischel 1974). The experimenter informs the child that he or she will leave the room and return with another of the chosen treats. The child is asked to choose whether to take the immediate, smaller reward (i.e., one marshmallow) or to wait for the delayed, larger reward (i.e., two marshmallows). Even if he or she decides to wait, the child must resist the temptation to take the smaller reward. The time the child is able to resist temptation and wait for the larger reward is taken as measure of DoG. After a maximum waiting time (typically 15–20 min in experiments with young children), the experimenter returns.

Because the ability to delay gratification increases with age, the maximum waiting time has to be extended as children get older (Green et al. 1994; Steinberg 2007). Many experiments with older children and adolescents have therefore adopted a dichotomous measure of DoG. In this procedure, two alternatives are presented: one smaller, immediate reward and one larger, delayed reward (e.g., in a month). Once the choice has been made, it cannot be changed (Silverman 2003).

It has previously been proposed that DoG measures for adults require not only meaningful delay intervals (days and weeks instead of minutes), but also meaningful und attractive rewards (Wulfert et al. 2002). However, it is difficult to find viable and non-trivial rewards for adults. Consequently, many researchers have used questionnaires to assess DoG in adults. For example, Ray and Najman (1986) used 12 questions with a yes/no answer format to tap DoG behaviors (e.g., "Are you good at saving your money rather than spending it straight away?"). Witt (1990) employed the same questionnaire, revised for use with a 5-point Likert scale. Ward, Perry, Woltz, and Doolin (1989) used a forced-choice response format to assess DoG in academic domains (e.g., "Go to a favorite concert and risk getting a bad grade, OR Stay home and study to get a better grade."). Likewise, the Academic Delay of Gratification Scale (ADOGS) developed by Bembenutty and Karabenick (1998) requires students to choose between two alternatives (e.g., "Study a little every day for an exam in this course and spend less time with your friends, OR Spend more time with your friends and cram just before the test."). A general problem of assessing DoG by self-report questionnaires is that responses may be affected by social desirability bias.

A different approach requires respondents to choose between two hypothetical money rewards (e.g., Madden et al. 1997). Although this technique refers to behavior, it does not measure observable behavior in the same way as the original DoG paradigm. A further development of this approach is the delay discounting (DD) paradigm, which investigates how much the value of a delayed hypothetical money reward decreases as the length of the delay increases (e.g., Critchfield and Kollins 2001).

All of the approaches to assessing DoG in adulthood described thus far depart from the original idea of a behavioral measure of self-control. However, two more rarely adopted approaches seem to measure DoG behavior in adults in a meaningful way. First, Funder and Block (1989) presented respondents with two real amounts of money. The disadvantage of this procedure is that it clearly becomes expensive as the number of trials increases. In a more economical variant, a single trial was applied to differentiate people with high vs. low ability to delay gratification (e.g., \$7 vs. \$10, Wulfert et al. 2002). Second, Knolle-Veentjer et al. (2008) offered snacks in a series of trials on a board game designed to be interesting to adults. The authors implemented a large number of trials (70 times 2 vs. 4 snacks such as chocolate drops, gummy bears, or crisps), leading to a decrease in impulsive choices over the course of the game.

Localization of delay of gratification in its nomological network

Beyond these problems of DoG assessment, much theoretical work has been done to localize DoG in the nomological network of the related constructs of self-control and self-regulation. Self-regulation, a complex multifaceted ability, involves internal and transactional processes that enable goal-directed activities to be maintained even under changing conditions (Baumeister and Heatherton 1996; Karoly 1993). Both self-regulation and DoG involve the active management of goals (Freund and Baltes 2002), and some authors see DoG as a measure of self-regulation (e.g., Mazur 1987). Likewise, Academic DoG and self-regulated learning strategies are strongly related (Bembenutty and Karabenick 2003). Successful self-regulation is necessarily accompanied by successful DoG—for example, when individuals prevent themselves from thinking about immediately available rewards (Mischel 1974).

Self-efficacy is a related, well-established construct that can be defined as an individual's confidence in his or her capacity to succeed in a specific situation (Bandura 1977). Self-efficacy plays a crucial role in the self-regulation of motivation (Bandura 1989, 1997) and can thus also be expected to be associated with DoG. For example, self-efficacy expectancies have been shown to predict persistence (i.e., DoG) with a fluid diet in hemodialysis patients (Rosenbaum and Ben-Ari Smira 1986).

Whereas self-regulation, self-control, and self-efficacy tend to be measured by self-report, delay discounting (DD) procedures involve a series of choices. DD can be

described as the degree to which the subjective value of a delayed reward decreases as a function of time to its delivery (Reynolds et al. 2002). Most assessments of DD involve monetary choice procedures (e.g., Kirby et al. 1999). Although DoG and DD are often used interchangeably, the two concepts have both similarities and differences (Mischel et al. 1988). One difference warrants emphasis: While DoG procedures require real behaviors of the participants, DD procedures involve only hypothetical choices.

Both DoG and DD are used as measures of impulsivity. According to the prominent discounting model of impulsiveness (Ainslie 1975), rates of DD are positively correlated with impulsiveness. In other words, impulsivity means choosing immediate rewards over later, larger rewards. Impulsivity can thus be viewed as the inability to wait for delayed rewards (Monterosso and Ainslie 1999).

DoG as a predictor of cognitive performance and wellbeing

Numerous studies have investigated how DoG relates to cognitive abilities and academic success. Preschoolers with high DoG have also been found to have higher attention control (Mischel 1974). DoG in childhood is seen as a predictor of various adolescent competencies. Individuals with higher DoG in childhood are more intelligent in adolescence than are children with low DoG (Mischel 1974). In a longitudinal study, moreover, Duckworth and Seligman (2005) found that self-discipline in grade 8 students (measured by questionnaires and a monetary choice task) was a significantly better predictor of academic performance than was IQ. In adulthood, high DoG is connected with intelligence, academic achievement, and need for achievement (Ayduk et al. 2000). Further, high Academic DoG (Bembenutty 1999) has been found to predict academic performance, motivation, help seeking, self-efficacy, and goal orientation (Mischel 1961; Mischel and Metzner 1962).

High DoG is also related to a lower risk of aggressive and delinquent behavior in adolescence (Krueger et al. 1996) and seems to have a protective effect in reducing drug abuse and increasing life satisfaction and self-worth (Rosenbaum and Ben-Ari Smira 1986). In other words, DoG shields individuals against behavior that eventually leads to lower wellbeing. Ayduk et al. (2000) also found protective effects of high DoG against the negative effects of rejection sensitivity that may reduce wellbeing in the long term.

The present study

The main aim of the present study was to develop and evaluate a delay of gratification test for adults (DoG-A). Two important conditions were set. First, the test should be a true test of behavior (i.e., real rewards should be used). Second, both the rewards and the delay intervals should be meaningful to adults. Whereas sweets are almost universally meaningful to children, their rewarding effect on adults differs. Different rewards assumed to be meaningful to a larger proportion of adults were therefore used: money, magazines, and snacks. Another reason for using different types of rewards was to investigate domain independence. The evidence suggests that DoG is a domain-specific construct (Bembenutty 1999; Ward et al. 1989)—that is, the ability to delay gratification varies from domain to domain.

A further aim of the study was to investigate how the DoG-A test relates to external variables. First, we assessed construct validity by calculating the correlations between DD, self-regulation, motivational self-concept, and relevant aspects of personality. Second, if the DoG-A captures a facet of self-regulation, it should be positively correlated with wellbeing and negatively correlated with psychopathological variables. Third, because the DoG-A is intended as a behavioral measure of self-regulation or self-motivation and not of cognitive ability, it should not be found to correlate significantly with cognitive function (or executive function, in particular).

Finally, we conducted regression analyses to identify the predictors of DoG-A and wellbeing. We hypothesized that DD, as a behavior-based measure, would be a stronger predictor of DoG than would self-regulation, motivational self-concept, or personality variables. Further, we investigated whether DoG-A was able to predict wellbeing over and above established self-report measures of motivational competence and health.

Method

Participants

A total of 147 adults, ranging in age from 60 to 94 years, participated in the study. The data of two participants were incomplete due to visual impairment. The data of eight further participants contained too many missing values or obvious outliers for meaningful analysis. As such, the present analyses were based on the data of 137 participants. All participants were cognitively healthy, community-dwelling individuals recruited from the greater Zurich area, Switzerland, via the University for Seniors (a weekly event for individuals aged 65 and older), old people's homes, and an advertisement in a magazine for seniors. Subjects participated voluntarily after receiving oral or written information about the study. The study population was stratified for age group (60–69, 70–79, 80+), sex, and education ($<$ vs. $\geq$13 years).

Procedure

Participants were administered a comprehensive questionnaire and several cognitive tests. The questionnaire contained the delay discounting test (DDT) and self-report measures of self-regulation, motivational self-concept, personality, and wellbeing. The DoG-A and the cognitive tests were administered during a 60–90 min session held either at the University or in the participant's home. The questionnaire was sent to participants via mail at least 1 week before the test session. Participants completed the questionnaire on their own, but were told that assistance was available if necessary.

The testing session, which was conducted by graduate students with training in psychological and neuropsychological assessment, took place in a comfortable room. The tasks were administered to all participants in the same order. Participants signed a consent form and were given feedback on their performance. In addition to travel expenses, participants received a magazine, snacks, and 10 Swiss Francs (approx. US$ 8.30) as compensation for their time.

Delay of gratification test for adults (DoG-A)

The DoG-A is a behavioral measure of motivational self-regulation. Four decision tasks involving four different types of rewards—snacks, hypothetical money, real money, and magazines (partly adapted from Knolle-Veentjer et al. 2008; Wulfert et al. 2002)—are embedded in a board game. To conceal the true aim of the test, the experimenter tells the participant that it is designed to measure their preferences and interests. Participant and experimenter take turns in moving a counter through the streets of a fictitious city. At each field on the board, the player draws a card and has to make a decision.

Eliciting of preferences

When contacting the participant by telephone to arrange a date for the testing session, the experimenter asks about the participant's snack preferences: "I'd like to offer you some nibbles at your meeting. Do you like chocolate? Or do you prefer salty snacks, such as cheese snacks?" The experimenter tries to elicit the participant's preferences to make the necessary preparations for the testing session. Appropriate snacks are also available for diabetic participants. This approach ensured that the participant was presented with an attractive incentive when asked to delay gratification.

Preparation

The testing session begins with the choice of two of five preselected snacks. The experimenter puts the snacks on the table and asks, "Which of these snacks would you prefer at the moment? Which would you go for right now? Is there a second snack you'd like at the moment?" The participant chooses two snacks. The experimenter then prepares the material on the table: the board, two counters, a dice, two packs of game cards, and the snacks on a plate.

Procedure

Next, the rules are explained. Participant and experimenter take turns tossing the dice and moving their counter through the streets of a fictitious city. At each field on the board, the player enters a shop and has to make a decision. There are two packs of cards: one for the participant and one for the experimenter. The cards drawn by participants pose questions about the products available in the shop. For example, the participant is asked, "There are black and red pullovers on sale. Do you like the black or the red pullover better?" The answer is recorded, but is not analyzed any further (these filler items conceal the true aim of the test). The fields on the board have one of seven colors. On each card, there are seven questions, one for each color. The participant is asked to read the one that corresponds to the color of the field the counter has landed on.

Whenever the experimenter draws a card, he or she offers the participant one of four different rewards:

1. Snacks: In eight trials, the participant has to decide between 1 piece of chocolate (or whatever the snack is) immediately and 2 pieces in 2 h (at the end of the session). The question is framed to be in line with the game: "Imagine you're in a café or a bakery and the sales assistant offers you…"
2. Hypothetical money: In further eight trials, two hypothetical money gifts are offered (immediately vs. in 1 month). The delayed amount is always CHF 10.00 and the immediate amount varies from CHF 6.00 to 9.50 in steps of .50, presented in the following order: 9.50, 6.00, 6.50, 9.00, 8.50, 7.00, 7.50, and 8.00. The question is, "Imagine that a friend of yours has won some money in the lottery. He or she wants to give you some money as a present. But you have to choose between CHF 6 now and CHF 10 in one month…"
3. Real money: The board has two fixed "event fields." When landing on one "event field," participants are offered real money. The experimenter puts CHF 8 in coins and a CHF 10 note on the table, and explains that the participant can choose between CHF 8 immediately or CHF 10 to be sent by mail in 1 month, together with the participant's individual feedback on the study results. If the counter does not directly land on this field, it nonetheless stops here. This procedure guarantees that this trial occurs at approximately the same time for each participant during the game.

4. Magazines: When landing on the other "event field," participants are offered magazines. The experimenter puts at least ten magazines from different categories (history, travel, women's interest, health, politics, gardening, dogs, etc.) on the table and says, "I would like to give you a magazine. Please choose one that interests you and that you would like to take home with you." After the participant has made his or her choice, the experimenter continues, "Which second magazine would you like to read and take home with you?" The final instruction is as follows: "You can take this magazine [the one chosen first] home with you. I would like to offer you both magazines, but in that case I would send you them by mail in one month, together with your individual feedback on the study results. Now you have to decide: Would you like one magazine now or two magazines in one month?"

The test ends when all of the experimenter's 18 cards have been played, i.e., when all 18 trials have been conducted. The order of the trials is S1, S2, H1, H2, S3, S4, H3, H4, S5, S6, H5, H6, S7, S8, H7, H8, with S denoting the snacks trials and H the hypothetical money trials. The real money trial occurs approximately after one third of the game, the magazine trial after two thirds.

Calculation of scores

A subscore is then calculated for each of the four types of reward. In the case of snacks and hypothetical money, the score is equal to the number of delayed rewards (range 0–8). In the case of real money and the magazines, the score is a dichotomous variable (0 = immediate, 1 = delayed reward). Two composite scores can also be calculated (see below for details).

Other variables

Delay discounting test

As a behavioral measure of self-control, we used the Swiss–German version (Forstmeier and Maercker 2010) of the Delay Discounting Test (DDT), also called the Monetary Choice Questionnaire (Kirby et al. 1999), the reliability (consistency) of which has been shown to be very high (Forstmeier and Maercker 2010). Participants were presented with a fixed set of 27 choices between smaller, immediate rewards and larger, delayed rewards. For example, participants were asked "Would you prefer CHF 68 today, or CHF 69 in 92 days?" The 27 items were grouped into three magnitude categories: small (CHF 32–44), medium (CHF 63–76), and large (CHF 95–107). Discounting rates were estimated on the basis of the pattern of 27 choices. They were first estimated separately for each magnitude category and then averaged as the geometric mean to calculate a global discounting rate. Discounting curves have been shown to be best described by a hyperbolic decay function (Mazur 1987):

$$V = \frac{A}{1 + kD} \quad (1)$$

where V is the present value of the delayed reward A at delay D, and k is a free parameter that determines the discounting rate k increases with the individual's preference for immediate rewards. Therefore, a higher discounting rate k can be interpreted as lower self-control or higher impulsiveness (Rachlin 1974). The validity of k as a behavioral measure of self-control/impulsiveness is indicated by its correlations with impulsiveness (Green and Myerson 2004; Kirby et al. 1999; Richards et al. 1999).

Self-regulation

Three aspects of self-regulation were assessed. Two five-item scales from the Volitional Components Questionnaire (VCQ, Kuhl and Fuhrmann 1998) assessed motivation regulation (e.g., "I can usually motivate myself quite well if my determination to persevere weakens.") and decision regulation (e.g., "When I think about doing or not doing something, I usually arrive at a decision quickly."). Participants rated their agreement with each statement on a 4-point scale. The alpha coefficient was .76 for motivation regulation and .71 for decision regulation. The locomotion scale of the Locomotion and Assessment Questionnaire (LAQ, Kruglanski et al. 2000) was used to measure activation regulation. The scale consists of 10 statements on activating oneself or commencing an action (e.g., "When I decide to do something, I can't wait to get started."). Participants rated their agreement with each item on a 6-point scale. The alpha coefficient was .64.

Motivational self-concept

Three aspects of the motivational self-concept were assessed. The General Self-Efficacy scale (GSE, Scholz et al. 2002) was used to assess the "broad and stable sense of personal competence to deal effectively with a variety of stressful situations" (Scholz et al. 2002, p. 243). Participants rated 10 items (e.g., "I am confident that I could deal efficiently with unexpected events.") on a 4-point scale. Dispositional optimism was assessed with the Life Orientation Test–Revised (LOT-R, Scheier et al. 1994). Finally, internal locus of control (IPC), or the generalized expectancy of being able to exert control over events, was assessed by the Internality scale of the Internality, Powerful Others, and Chance Scale (Levenson 1974).

Personality

All six facets of the conscientiousness scale (competence, order, dutifulness, achievement striving, self-discipline, deliberation) and the impulsiveness facet of the neuroticism scale from the NEO Personality Inventory–Revised (NEO-PI-R, Costa and McCrae 1992) were administered. Each facet comprises eight items and is rated on a 5-point scale. Internal consistencies were between $\alpha = .48$ and .74.

Wellbeing

Satisfaction with life was assessed with the Satisfaction With Life Scale (SWLS, Diener et al. 1985), a five-item measure of overall life satisfaction, in which higher scores indicate greater life satisfaction. General affectivity was assessed by the Positive and Negative Affect Schedule (PANAS, Watson et al. 1988)—a list of 20 adjectives, each of which is rated on a 5-point scale. Of the 20 adjectives, 10 form a positive affect scale (e.g., enthusiastic, excited) and 10 form a negative affect scale (e.g., upset, afraid). Participants were asked to think about how they were feeling in general. Ratings were averaged to generate positive and negative affect scores. Depressive symptoms were assessed with the Geriatric Depression Scale (GDS, Yesavage et al. 1983), which is widely used in geriatric research. We used the 15-item version with a yes/no answer format. In addition, we used three subscales of the Brief Symptom Inventory (BSI, Derogatis 1993) to assess anxiety, hostility, and somatization. Participants rated their agreement with each item on a 5-point scale. Finally, the 10-item Perceived Stress Scale (PSS, Cohen et al. 1983), which is widely used in research, measures the perceived stress of the current life situation. Participants rated the frequency of unpredictable, uncontrollable, and overwhelming events in the past month on a 5-point scale.

Cognitive function

Current cognitive status was assessed by eight cognitive performance tests. Memory was assessed using the "Word List Recall" subtest of the neuropsychological battery developed by the Consortium to Establish a Registry for Alzheimer's Disease (CERAD, (Morris et al. 1989). Specifically, participants tried to recall ten words they had learned 10–15 min previously. Verbal Fluency was assessed with the Animal Naming Task (Morris et al. 1989): participants were asked to name as many different animals as possible in 60 s. A German vocabulary test (Wortschatztest, WST, Schmidt and Metzler 1992) was used to assess crystallized intelligence (or knowledge). The test consists of 42 lines of six words. One of the words in each line is real; five are nonsense. Participants are asked to identify the real word in each line. Difficulty increases from line to line. The Digit Symbol Substitution Test (DSST) of the Wechsler Adult Intelligence Scale-III (WAIS-III, Wechsler 1997) was employed to assess visual-motor speed. Participants were asked to match as many digits as possible to a set of corresponding symbols in 120 s. Trail Making Tests A and B (Reitan 1958) were used to measure cognitive speed (TMA-A) and task switching (TMT-B). In these tests, participants connect numbers in ascending order (A) or alternating numbers and letters in ascending order (B). The latter test is interpreted as a measure of executive function. Two further executive tests were employed: the Digit Span Backward and the Stroop Color-Word Test. The Digit Span Backward, a measure of working memory, is a WAIS-III subtest (Wechsler 1997). The Stroop Color-Word Test (SCWT, Stroop 1935) measured inhibition of prepotent responses. Participants were asked to read aloud or name the stimuli on each card (color names on card 1, color of the patches on card 2, and color of the ink on card 3) as quickly as possible without making mistakes. We calculated the ability to inhibit prepotent responses by subtracting the time needed to respond to card 2 from the time needed to respond to card 3.

Education

Participants were asked to state their highest level of education. Individuals with a university (master's) degree were coded as having completed 18 years of education, and participants with a PhD or MD as having completed 21 years of education, regardless of their actual number of years in school.

Statistical analyses

DoG-A subscores were calculated as described above. To calculate a composite DoG score, we first dichotomized the two continuous variables (snacks and hypothetical money), with the scale mid-point as cut-off (0–4 vs. 5–8). Two composite scores were constructed, one the sum of all four subscores and one the sum of the Snacks, Hypothetical money, and Real money subscores.

Bivariate correlations were calculated to explore relationships among all variables. In a first step, we have tested all bivariate correlations on the 5% alpha-level. In a second step, in order to adjust for multiple testing, we tested the correlations using a Bonferroni-corrected alpha-level (i.e., alpha = .003 for the intercorrelations in Table 3 alpha = .0015 for the correlations with external variables in Table 4).

Two hierarchical multiple regression analyses were then calculated to identify the predictors of DoG-A and

Table 1 Overview over the measures used and allocation to the composite scores

Constructs	Measures	Composite scores		
		Mot.	Wellb.	Cog.
Delay discounting rate (*k*)	Delay Discounting Test (DDT)			
Self-regulation				
Motivation regulation	Scale Motivation Regulation of the Volitional Components Questionnaire (VCQ)	x		
Decision regulation	Scale Decision Regulation of the Volitional Components Questionnaire (VCQ)	x		
Activation regulation	Scale Locomotion of the Locomotion and Assessment Questionnaire (LAQ)	x		
Motivational self-concept				
Optimism	Life Orientation Test–Revised (LOT-R)	x		
General self-efficacy	General Self-Efficacy scale (GSE)	x		
Internal locus of control	Internality scale of the Internality, Powerful Others, and Chance Scale (IPC)	x		
Personality				
Conscientiousness	Conscientiousness scale of the NEO Personality Inventory–Revised (NEO-PI-R)			
Impulsiveness	Impulsiveness facet of the neuroticism scale of the NEO-PI-R			
Current psychological wellbeing				
Satisfaction with life	Satisfaction With Life Scale (SWLS)		x	
Positive affect	Positive affect scale of the Positive and Negative Affect Schedule (PANAS)		x	
Negative affect	Negative affect scale of the Positive and Negative Affect Schedule (PANAS)		x	
Depressive symptoms	Geriatric Depression Scale (GDS)		x	
Anxiety	Anxiety scale of the Brief Symptom Inventory (BSI)		x	
Hostility	Hostility scale of the Brief Symptom Inventory (BSI)		x	
Somatization	Somatization scale of the Brief Symptom Inventory (BSI)			
Perceived stress	Perceived Stress Scale (PSS)			
Cognitive function				
Memory	Word List Recall subtest of the Consortium to Establish A Registry for Alzheimer's Disease (CERAD) battery			x
Verbal fluency	Animal Naming Task subtest of the CERAD			x
Verbal intelligence	Wortschatztest (WST) (German Vocabulary Test)			x
Visual-motor speed	Digit Symbol Substitution Test (DSST) of the Wechsler Adult Intelligence Scale-III (WAIS-III)			x
Cognitive speed	Trail Making Test A (TMT-A)			x
Task switching	Trail Making Test B (TMT-B)			x
Working memory	Digit Span Backward of the WAIS-III			x
Inhibition of prepotent responses	Stroop Color-Word Test (SCWT)			x

Mot., motivational competence; Wellb., psychological wellbeing; Cog., cognitive function

wellbeing, respectively. In the regression analysis predicting DoG, sociodemographic variables (age, sex, and education) were entered in step 1, health variables (psychological wellbeing, somatic complaints, and perceived stress) in step 2, and motivational variables (delay discounting, motivation regulation, optimism, personality factors) in step 3. In the regression analysis predicting psychological wellbeing, sociodemographic variables (age, sex, education, being married, living with others) were entered in step 1, somatic complaints in step 2, cognitive function in step 3, perceived stress in step 4, self-reported motivational competence in step 5, and DoG in step 6. The DoG-A Snacks subscale was used as continuous scale.

To reduce the number of variables and to minimize floor and ceiling artifacts and other forms of measurement error, we calculated three composite measures (psychological wellbeing, cognitive function, and self-reported motivational competence) by converting the component tests to *z* scores, using the baseline mean and standard deviation of all study participants, and averaging the *z* scores (see Table 1). The composite score for psychological wellbeing comprised satisfaction with life (SWLS), positive and negative affect (PANAS), depression (GDS), anxiety and hostility (BSI). The composite score for cognitive function comprised the eight cognitive performance tests described above. Finally, the composite score for self-reported

motivational competence comprised the three self-regulation and the three motivational self-concept scales described above.

Results

Sample characteristics

Characteristics of the sample and descriptive data at baseline are given in Table 2. The 137 participants (62% women) had a mean age of 74 years (age range 60–94) and a mean of 13.4 years' education. Almost half lived with a partner or other persons, 29% lived alone, and 22% in old people's homes. On average, women were less educated and more likely to live alone than men.

Delay of gratification score

Snacks subscore

The mean number of delayed rewards on the eight trials was 5.6 (SD 2.3; see Table 2). In other words, the mean percentage of trials in which the delayed reward was chosen was 70.5%. There was no sex difference (t (135) = −1.4; p = .178). As shown in Fig. 1a, the participants' decisions across the eight trials generate a saturation curve: 50% of participants chose delayed reward in the first trial, rising to 76% in the final trial. The best fitting curve is an exponential function:

$$y = 76\left(1 - e^{-1x}\right) \quad (2)$$

where x is the trial number and y is the estimated value. With increasing trial number, the slopes of this exponential function decrease, starting with 28.0 at the first trial and dropping to .03 at the final trial. Estimated values and slopes are reported in the note to Fig. 1a. In view of this saturation curve, we decided to calculate a mean score for the number of delayed rewards weighted by the slope of the curve at the respective trial. As a result, the first trials are weighted much more heavily than the last ones, leading to M = 25.1 (SD = 17.7). Using this weighted score in the following analyses did not lead to appreciably different results than using the unweighted score. Therefore, the latter were used in all analyses reported below.

Hypothetical money subscore

The mean number of delayed rewards on the eight trials was 3.4 (SD 2.7; see Table 2). In other words, the mean percentage of trials in which the delayed reward was chosen was 42.5%. There was no sex difference (t (135) = .9;

Table 2 Baseline characteristics of the sample by sex (N = 137)

Characteristic	Total	Sex		
		Male	Female	t/χ^2 (df = 135)
Age (years), M (SD)	73.9 (8.1)	74.2 (7.1)	73.8 (8.6)	.28
Sex (% female)	62.0			
Education (years), M (SD)	13.4 (2.5)	14.3 (2.8)	12.9 (2.2)	3.1*
Living situation				30.3*
Single, at home (%)	29.2	7.7	42.4	
With partner (%)	46.7	75.0	29.4	
With family member (%)	1.5	.0	2.4	
Old people's home (%)	21.7	17.3	24.7	
With other persons (%)	.7	.0	1.2	
Delay of gratification for adults (DoG-A)				
Snacks, M (SD)[a]	5.6 (2.3)	5.3 (2.5)	5.9 (2.1)	−1.4
Hypothetical money, M (SD)[a]	3.4 (2.7)	3.6 (2.8)	3.2 (2.7)	.9
Real money (% delayed reward)	65.0	71.2	61.2	1.4
Magazines (% delayed reward)	59.1	71.2	51.8	5.0*
Composite score A (all subscales), M (SD)[b]	2.4 (1.3)	2.6 (1.1)	2.2 (1.1)	1.7
Composite score B (without mags), M (SD)[c]	1.8 (1.0)	1.9 (.9)	1.7 (.9)	.8

[a] DoG-A subscales snacks and hypothetical money have a value range of 0–8

[b] Composite score A has a range of 0–4

[c] Composite score B has a range of 0–3

* $p < .05$

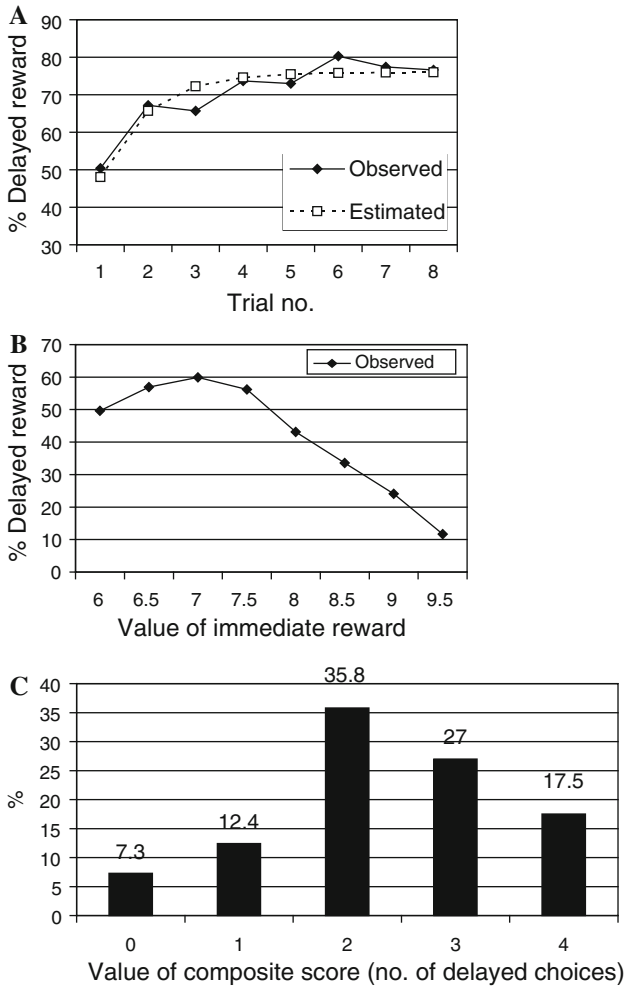

Fig. 1 Percentage of participants choosing the delayed reward. **a** Eight trials of the snacks subscale. The *continuous line* shows the observed values (50.4, 67.2, 65.7, 73.7, 73.0, 80.3, 77.4, and 76.6). The *dotted line* shows the estimated values using the exponential function $y = 76\ (1 - e^{-1x})$ (48.0, 65.7, 72.2, 74.6, 75.5, 75.8, 75.9, and 76.0). The slopes of the exponential function are 28.0, 10.3, 3.8, 1.4, .5, .2, .07, and .03, for the eight items. **b** Eight trials of the Hypothetical money subscale. The delayed reward was always CHF 10, the value of the immediate reward ranged from CHF 6.00 to 9.50. The *line* shows the observed values (49.6, 56.9, 59.9, 56.2, 43.1, 33.6, 24.1, and 11.7). **c** Five possible values of the composite score A (0 = delayer in no subscale, 4 = delayer in all subscales)

$p = .378$). Whereas the immediate rewards remained equal in the snacks trials, they varied from CHF 6.00 to CHF 9.50 in the hypothetical money trials (the delayed reward remained stable at CHF 10.00). As shown in Fig. 1B, the number of participants choosing the delayed reward decreased as the amount of the immediate reward increased. The percentage of participants choosing the delayed reward of CHF 10.00 started at 50% when CHF 6.00 was offered as an immediate reward, rising to between 50% and 60% for an immediate reward of CHF 6.50, 7.00, or 7.50, and dropping below 50% for an immediate reward of between CHF 8.00 and 9.50.

Real money and magazines subscores

For the two dichotomous items, the percentage of participants choosing the delayed reward was 65.0% for real money and 59.1% for magazines. There was no sex difference for real money (χ^2 (1) = 1.4, $p = .235$), but more men than women chose the delayed reward in the case of magazines (71.2% vs. 51.8%; χ^2 (1) = 5.0, $p = .025$).

Intercorrelations of subscores

The intercorrelations of the subscales were low to medium (Table 3). The correlation between Snacks and Hypothetical money was $r = .27$ ($p = .001$); that between Real money and Hypothetical money was $r = .36$ ($p < .001$). The Magazines subscale did not correlate with any of the other three subscales.

Composite score

To calculate a composite DoG score, we first dichotomized the two continuous variables, using the scale mid-point as cut-off values. These values separated delayers from non-delayers with a distribution approximately 1:2 (73.7% delayers in the Snacks scale; 37.2% delayers in the Hypothetical money scale). Two composite scores were constructed. Composite score A is the sum of all four subscores (0 = non-delayer in all subscales, 4 = delayer in all subscales). Because Magazines did not correlate with any of the other subscales, composite score B was additionally calculated as the sum of the Snacks, Hypothetical money, and Real money subscores (range 0–3). Figure 1c shows the frequency distribution of composite score A, which approached normal distribution. Mean composite score A was 2.4 (SD = 1.3); mean composite score B was 1.8 (SD = 1.0). Internal consistency (Cronbach's alpha) was .39 for composite score A and .41 for composite score B.

Correlations between DoG and external variables

Criterion validity

To evaluate the criterion validity of the DoG-A, we calculated its bivariate correlations with measures of DD, self-regulation, motivational self-concept, and relevant aspects of personality (see Table 4). The correlations of the DoG-A subscores and composite scores with the DD rate were the highest and most consistent. Correlations with the general DD rate ranged from $r = -.22$, $p = .010$ (Magazines) to $r = -.46$, $p < .001$ (Hypothetical money) and $r = -.45$, $p < .001$ (composite score A). Note that a high DD rate indicates low self-control. The better the respondent is able

Table 3 Intercorrelations of the DoG-A subscales ($N = 137$)

Scales	Subscales			Composite scores	
	Hypoth. money	Real money	Mags	A (all subscales)	B (without mags)
Subscales					
Snacks	.27*	.13	.05	.51*	.59*
Hypothetical money		.36*	.06	.60*	.68*
Real money			.11	.65*	.72*
Magazines				.55*	.13
Composite scores					
Composite Score A (all subscales)					.90*
Composite Score B (without mags)					

The values represent Pearson correlations (between two continuous variables), point-biserial correlations (between a continuous and a dichotomous variable), or phi coefficients (between two dichotomous variables)

* $p < .05$. To adjust for 15 tests, the critical alpha-level is reduced to .003

to delay gratification, the lower the rate at which he or she discounts future rewards.

Of the four DoG-A subscales, Snacks showed the most consistent pattern of results, correlating with self-reported motivation regulation ($r = .30$, $p < .001$), optimism ($r = .17$, $p = .044$), and two facets of Conscientiousness: dutifulness ($r = .19$, $p = .030$) and deliberation ($r = .22$, $p = .011$). Motivation regulation and optimism also correlated with at least one composite score (see Table 4).

After adjusting alpha for multiple testing, only the correlations of DoG with DD and DoG Snacks with motivation regulation remained significant.

Correlation with cognitive function

As a behavioral measure of self-regulation or self-motivation, the DoG-A was not expected to correlate significantly with cognitive function (or executive function, in particular). The pattern of correlations is in line with this expectation (see Table 4).

Correlation with psychological wellbeing

If the DoG-A captures a facet of self-regulation, it should be positively correlated with wellbeing and negatively correlated with psychopathological variables. As shown in Table 4, only the Snacks subscale correlated with satisfaction with life ($r = .31$, $p < .001$), negative affect ($r = -.20$, $p = .019$), depressive symptoms ($r = -.25$, $p = .003$), anxiety ($r = -.29$, $p = .001$), hostility ($r = -.20$, $p = .017$), and perceived stress ($r = -.24$, $p = .006$). Composite scores A and B correlated with satisfaction with life, depressive symptoms, anxiety, and hostility (only B). After adjusting alpha for multiple testing, only the correlations of DoG Snacks with satisfaction with life and anxiety remained significant.

Predicting delay of gratification

Given the findings that the DD rate as well as self-reported motivation regulation, optimism, and personality factors (dutifulness and deliberation) are associated with DoG-A Snacks scale, the question arises which of these constructs is best able to predict the Snacks score. The results of hierarchical multiple regression analyses are summarized in Table 5. When sociodemographic and health variables were controlled, the DD rate was the best predictor of DoG Snacks ($\beta = -.27$, $p = .001$). Although motivation regulation remained a significant predictor of DoG Snacks ($\beta = .17$, $p = .050$), optimism, dutifulness, and deliberation lost their predictive power. These predictors explained 24% of the variance (corrected $R^2 = .24$; effect size $\varepsilon^2 = .32$; $F(135) = 4.86$; $p < .001$). The last column for each model in Table 5 shows how much additional variance was explained when the respective block was included in the final step of the analysis. Motivational variables (i.e., DD, motivation regulation) explained 13% of the variance in DoG Snacks when all other variables were controlled.

We also conducted a hierarchical multiple regression analysis predicting composite score A (see Table 5). DD rate was the only variable that predicted delay of gratification ($\beta = -.46$, $p < .001$). Neither health nor other motivational variables showed a significant β. In this model, 26% of the variance was explained (corrected $R^2 = .26$; effect size $\varepsilon^2 = .35$; $F(135) = 5.37$; $p < .001$).

Predicting psychological wellbeing

The bivariate correlations showed small to medium-sized associations between DoG-A Snacks and most psychological wellbeing scales (Table 4). We conducted a further hierarchical multiple regression analysis to investigate

Table 4 Correlations between DoG-A and delay discounting, self-regulation, motivational self-concept, personality, psychological wellbeing, and cognitive function ($N = 137$)

Variables	*M* (SD)	Snacks	Hypoth. money	Real money	Mags	Composite A (all)	Composite B (without mags)
Delay discounting rate (*k*): general	.027 (.06)	−.31*	−.46*	−.34*	−.22*	−.45*	−.43*
Small reward size	.036 (.07)	−.33*	−.45*	−.30*	−.18*	−.44*	−.42*
Medium reward size	.030 (.07)	−.29	−.41*	−.34*	−.21*	−.43*	−.41*
Large reward size	.024 (.06)	−.27	−.44*	−.33*	−.24*	−.44*	−.39*
Self-regulation, *M* (SD)							
Motivation regulation (VCQ)	8.4 (2.5)	.30*	.16	.02	−.14	.11	.20*
Decision regulation (VCQ)	10.3 (2.4)	.002	.02	−.10	−.05	−.06	−.04
Activation regulation (LAQ)	47.7 (6.5)	−.08	−.09	−.19*	.002	−.09	−.11
Motivational self-concept, *M* (SD)							
Optimism (LOT-R)	16.4 (3.4)	.17*	.12	.13	.10	.24*	.24*
General self-efficacy (GSE)	28.3 (4.3)	.04	.04	.05	.10	.13	.10
Internal locus of control (IPC)	37.1 (4.6)	.08	−.05	−.06	.13	.07	.01
Personality (NEO-PI-R), *M* (SD)							
Conscientiousness-competence	22.9 (3.6)	.03	.04	.07	.01	.08	.09
Conscientiousness-order	20.6 (3.9)	−.08	−.05	−.17*	−.07	−.11*	−.10
Conscientiousness-dutifulness	26.0 (3.4)	.19*	−.05	.03	.14	.15	.10
Conscientiousness-achievement striving	19.6 (4.5)	.03	.11	.03	.14	.13	.08
Conscientiousness-self-discipline	22.4 (4.2)	−.05	−.11	−.18*	−.07	−.17*	−.16
Conscientiousness-deliberation	19.0 (4.3)	.22*	.10	.06	.03	.14	.15
Conscientiousness total score	130.7 (17.0)	.06	.002	−.05	.03	.02	.01
Neuroticism-impulsiveness	12.5 (3.4)	−.11	.003	−.01	.13	.03	−.03
Current psychological wellbeing							
Satisfaction with life (SWLS)	27.2 (4.3)	.31*	.11	.09	.13	.20*	.17*
Positive affect (PANAS)	35.2 (5.3)	.08	−.06	−.15	.01	−.05	−.06
Negative affect (PANAS)	16.8 (5.4)	−.20*	.01	−.06	−.05	−.08	−.07
Depressive symptoms (GDS)	1.8 (2.1)	−.25*	−.10	−.02	−.09	−.18*	−.17*
Anxiety (BSI)	2.2 (2.6)	−.29*	−.09	−.10	−.09	−.20*	−.19*
Hostility (BSI)	1.4 (1.8)	−.20*	−.03	−.09	.02	−.14	−.17*
Somatization (BSI)	3.0 (3.0)	−.01	−.08	−.01	.00	−.03	−.03
Perceived Stress (PSS)	11.5 (5.8)	−.24*	−.02	.05	.00	−.07	−.08
Cognitive function							
Word List Recall (CERAD)	7.2 (2.2)	−.01	−.12	.03	.05	−.01	−.02
Verbal fluency (CERAD)	22.2 (6.4)	−.02	−.07	.05	.001	.04	.04
Verbal intelligence (WST)	34.1 (4.3)	.08	.002	−.01	.03	.06	.06
Digit symbol substitution test	51.1 (14.8)	−.02	−.21*	−.03	.11	−.01	−.07
Trail making test A	46.6 (20.2)	−.04	.11	.05	−.002	−.01	−.01
Trail making test B	115.1 (51.6)	−.03	.01	−.02	−.10	−.09	−.05
Digit span backward	6.0 (1.8)	.06	−.02	.17	.03	.14	.15
Stroop color-word test	23.1 (14.6)	−.10	.06	.02	−.06	−.07	−.05

The values represent Pearson correlations (between two continuous variables) or point-biserial correlations (between a continuous and a dichotomous variable)

VCQ, Volitional Components Questionnaire; LAQ, Locomotion and Assessment Questionnaire (locomotion scale); LOT-R, Life Orientation Test–Revised; GSE, General Self-Efficacy scale; IPC, Internality, Powerful Others, and Chance Scale; NEO-PI-R, NEO Personality Inventory–Revised; SWLS = Satisfaction With Life Scale; PANAS, Positive and Negative Affect Schedule; GDS, Geriatric Depression Scale; BSI, Brief Symptom Inventory; PSS, Perceived Stress Scale; CERAD, Consortium to Establish A Registry for Alzheimer's Disease; WST, Wortschatztest (German Vocabulary Test)

* $p < .05$. To adjust for 34 tests, the critical alpha-level is reduced to .0015

Table 5 Summary of multiple hierarchical regression analyses predicting delay of gratification ($N = 137$)

	Predicting DoG-A snacks				Predicting DoG-A composite A			
	B	SE	β	ΔR^2	*B*	SE	β	ΔR^2
Step 1: Sociodemographic variables				.07*				.03
Age	−.06	.02	−.21*		−.02	.01	−.15	
Sex (1 = m; 2 = f)	.67	.38	.14		−.28	.19	−.12	
Education (in years)	−.09	.07	−.10		.01	.04	−.02	
Step 2: Health variables				.07*				.02
Psychological wellbeing[a]	1.24	.43	.35*		.18	.22	.12	
Somatic complaints (BSI)	.15	.07	.20*		.01	.03	.08	
Perceived Stress (PSS)	−.01	.04	−.02		.04	.02	.18	
Step 3: Motivational variables				.13*				.22*
Delay Discounting Rate (DDQ)	−.35	.11	−.27*		−.31	.05	−.46*	
Motivation regulation (VCQ)	.16	.08	.17*		−.01	.04	−.02	
Optimism (LOT-R)	−.06	.06	−.09		.05	.03	.18	
Dutifulness (NEO-PI-R)	−.02	.06	−.03		.03	.03	.16	
Deliberation (NEO-PI-R)	.07	.05	.14		.02	.02	−.02	

ΔR^2 in the case that the respective block is included in the final step of the analysis

BSI, Brief Symptom Inventory; PSS, Perceived Stress Scale; DDQ, Delay Discounting Questionnaire; VCQ, Volitional Components Questionnaire; LOT-R, Life Orientation Test-Revised; NEO-PI-R, NEO Personality Inventory-Revised

[a] Composite score of Satisfaction With Life Scale (SWLS), Positive and Negative Affect Schedule (PANAS), Geriatric Depression Scale (GDS), and the Anxiety and Hostility subscales of BSI

* $p < .05$

whether this association persisted when other factors influencing psychological wellbeing were controlled. In particular, we were interested in whether DoG-A as a behavioral measure was able to predict wellbeing over and above established self-report measures of motivational competence.

Table 6 summarizes two hierarchical regression analyses, one including DoG-A Snacks in the final step, the other including composite score A in the final step. The results of the two analyses are very similar (the values presented in the following refer to the Snacks score). The only sociodemographic variable predicting psychological wellbeing was age ($\beta = .16$, $p = .043$). Not surprisingly, somatic complaints predicted wellbeing in this sample of older people ($\beta = -.21$, $p = .002$). Cognitive function—measured as mean of all cognitive tests after z transformation—was not a significant predictor of wellbeing ($\beta = .10$, $p = .222$). The most important predictor of wellbeing was the perceived stress level ($\beta = -.45$, $p < .001$). Self-reported motivational competence—i.e., the mean of all self-regulation and motivational self-concept variables after z transformation—significantly predicted wellbeing ($\beta = .22$, $p = .001$). Finally, DoG-A was included in the final step of the model and still predicted wellbeing ($\beta = .21$, $p = .001$). This set of predictors explained 56% of the variance when DoG-A Snacks was entered in the final step (corrected $R^2 = .56$; effect size $\varepsilon^2 = 1.27$; $F(136) = 18.53$; $p < .001$), and 54% of the variance when the DoG-A composite score A was entered in the final step (corrected $R^2 = .55$; effect size $\varepsilon^2 = 1.22$; $F(136) = 17.32$; $p < .001$).

Discussion

The main aim of this study was to evaluate a newly developed delay of gratification test for adults (DoG-A). As expected, the findings indicated relative domain independence of the four reward types, as indicated by low-to-medium correlations between the four subscales and moderate internal consistency of the composite scores ($\alpha = .4$). The nomological net established by investigating the relations of the DoG-A with other constructs proved to be fairly meaningful. All four subscales showed the strongest and most consistent correlations with the delay discounting rate ($r = -.22$ to $-.46$). The Snacks subscale showed the most consistent pattern of results: moderate positive correlations with self-reported motivation regulation, optimism, dutifulness, and deliberation. Hierarchical regression analyses controlling for sociodemographic and health variables revealed that delay discounting was a better predictor of DoG Snacks than were self-report measures of self-regulation. Further, only the Snacks subscale correlated consistently with measures of psychological wellbeing.

Table 6 Summary of multiple hierarchical regression analyses predicting psychological wellbeing ($N = 137$)

	Step 6: DoG-A snacks				Step 6: DoG-A composite A			
	B	SE	β	ΔR^2	*B*	SE	β	ΔR^2
Step 1: Sociodemographic variables				.02				.02
Age	.01	.01	.16*		.01	.01	.16*	
Sex (1 = m; 2 = f)	.01	.09	.01		.09	.09	.07	
Education (in years)	−.01	.02	−.03		−.01	.02	−.04	
Married (0 = no; 1 = yes)	.07	.11	.06		.14	.12	.11	
Living with others (0 = no; 1 = yes)	−.03	.12	−.02		−.05	.13	−.03	
Step 2: Somatic complaints (BSI)	−.04	.01	−.21*	.03*	−.04	.01	−.19*	.03*
Step 3: Cognitive function[a]	.09	.07	.10	.01	.09	.08	.09	.01
Step 4: Perceived Stress (PSS)	−.05	.01	−.45*	.12*	−.06	.01	−.50*	.16*
Step 5: Self-reported motivational competence[b]	.21	.06	.22*	.04*	.21	.06	.22*	.04*
Step 6: Delay of gratification	.06	.02	.21*	.04*	.08	.03	.16*	.02*

ΔR^2 in the case that the respective block is included in the final step of the analysis. Psychological wellbeing is measured as composite score of Satisfaction With Life Scale (SWLS), Positive and Negative Affect Schedule (PANAS), Geriatric Depression Scale (GDS), and the Anxiety and Hostility subscales of BSI

BSI, brief symptom inventory; PSS, perceived stress scale

[a] Composite score of the CERAD subtests Word List Recall and Verbal Fluency, verbal intelligence (WST), Digit Symbol Substitution Test, Trail Making Test A and B, Digit Span Backward, and Stroop Color-Word Test

[b] Composite score of the Motivation Regulation and Decision Regulation subscales of the Volitional Components Questionnaire (VCQ), Locomotion subscale of the Locomotion and Assessment Questionnaire (LAQ), Life Orientation Test–Revised (LOT-R), General Self-Efficacy scale (GSE), Internality, Powerful Others, and Chance Scale (IPC)

* $p < .05$

Hierarchical regression analyses showed that DoG Snacks remained a significant predictor of wellbeing when sociodemographic variables, somatic complaints, cognitive function, perceived stress, and self-reported motivational competence were controlled.

Domain independence of reward types

The finding that the four DoG-A subscales showed low-to-medium intercorrelations and that the composite scores had moderate internal consistency ($\alpha = .4$) is not new. Funder, Block, and Block (1983) reported an even lower internal consistency ($\alpha = .3$) for their composite DoG score. Other research has applied factor analysis to identify further domains of DoG. For example, Ward et al. (1989) focused on the sociopolitical and achievement specificity of delay of gratification. The results of their factor analysis revealed several dimensions of DoG, namely, sociopolitical, career objective, and academic achievement. Related research on delay discounting has also found low correlations between domains. For example, Chapman (1996) reported a correlation of $r = .11$ between discounting rates in the domains of money and health.

Taken together, delay of gratification seems to be highly dependent on the situation (Metcalfe and Mischel 1999). The ability to delay gratification can be assumed to be affected by individuals' ongoing histories of experience with environmental contingencies, which may differ across domains. Thus, a global conceptualization of DoG seems to be of limited value.

In terms of the reliability of the DoG-A test, research is needed to investigate its test–retest reliability. Findings on the test–retest reliability of delay discounting measures (e.g., $r = .71$ after 1 year; Kirby 2009) point to relative stability over time. The stability of the DoG-A can be expected to be of similar magnitude, but data on its test–retest reliability are still pending.

Behavioral and self-report measures of self-regulation

The strongest (significant) correlations of DoG-A subscales with external variables were found with the delay discounting rate. Individuals with high DoG had a low discounting rate, indicating that both measures capture a similar aspect of self-control. Clearly, the high correlation of the DDT with the DoG-A Hypothetical money subscale is rooted in the fact that both concern the same domain.

Only four correlations of DoG-A Snacks and composite scores with self-reported self-regulation, motivational self-concept, and personality aspects were significant: motivation regulation, optimism, dutifulness, and deliberation. It makes sense that the strongest correlation was with motivation regulation: the items of this scale tap the ability to persevere by motivating oneself, which is vital for delaying

gratification. In contrast, decision and activation regulation imply swift action, which is almost the opposite of delaying gratification.

Although some correlations of DoG-A with self-concept and personality aspects were significant, they were rather low. This pattern of results is consistent with data collected in samples of children. For example, Krueger et al. (1996) found significant correlations of between $r = -.18$ and .16 with personality characteristics such as responsible, productive, and self-controlled. Funder et al. (1983) reported correlations of between $r = .2$ and .4 with similar personality characteristics. These low correlations may be attributable to behavioral and self-report measures capturing *different facets* of related constructs. This phenomenon is also found in other areas of psychological measurement; for example, implicit and explicit measures of motives are relatively independent (Schultheiss et al. 2009). An alternative explanation is that self-report measures are more prone to social desirability response bias. Because behavioral and self-report measures of self-control to a certain extent tap different aspects of self-control, it is important to use both approaches in research to ensure the comprehensive assessment of the construct.

Predicting psychological wellbeing

Although we did not investigate clinical populations in this study, we found meaningful correlations of the DoG-A Snacks subscore with depressive symptoms, anxiety, hostility, negative affect, satisfaction with life, and perceived stress. In addition, hierarchical regression analysis showed that DoG Snacks as well as the composite score remained a significant predictor of wellbeing when all other predictors were controlled.

There is a long tradition of research into self-control failure as basis for various psychopathological syndromes. Studies using self-report measures of self-control have shown meaningful correlations of self-control with psychological wellbeing, health behavior, and social skills (Kruglanski et al. 2000; Tangney et al. 2004). Most research on psychopathology applying DoG tasks has involved samples of children or adolescents. For example, high DoG has been shown to be associated with lower risk of aggressive and delinquent behavior in adolescence (Krueger et al. 1996), lower risk of drug abuse, higher life satisfaction, higher self-worth (Rosenbaum and Ben-Ari Smira 1986), lower risk of involvement in cigarettes, alcohol, and marijuana, and higher self-esteem (Wulfert et al. 2002). Research applying delay discounting procedures has found higher discounting rates (i.e., lower self-control) in people abusing alcohol and drugs (de Wit 2009), pathological gamblers (Reynolds 2006), and individuals with attention deficit hyperactivity disorder (Scheres et al. 2008), antisocial personality disorder (Petry 2002), obesity (Weller et al. 2008), schizophrenia (Heerey et al. 2007), and social anxiety (Rounds et al. 2007).

Why does only the DoG-A Snacks subscale correlate significantly with measures of wellbeing? One explanation is that the domain of food is more sensitive to self-control failure. Indeed, a study comparing the domains of food, alcohol, and money found reduced self-control with regard to food and alcohol, but not money (Odum and Rainaud 2003). In addition, money and magazines may be less important to older Swiss people than snacks because of the relative wealth of this age group in Switzerland. In other words, CHF 8 vs. 10 is too small a sum to be a powerful incentive for this group. According to Mischel (1974), "it is necessary to consider the determinants of the individual's choice to delay for the sake of more preferred delayed outcomes" (p. 287). Mischel expected DoG to be determined primarily by the value of the reward and the expectancy of success. Similarly, expectancy-value theory highlights expectancy and value as two major determinants of task choice (Feather 1990). The Snacks subscale may have had the highest value for older Swiss people.

Limitations

Several limitations of this study must be considered. First, given its cross-sectional design, the findings represent associations and not causal effects of delay of gratification. Longitudinal studies are needed to address further-reaching questions, such as how ability to delay gratification relates to emotional wellbeing or functional abilities, and to investigate test–retest reliability and stability versus change over time. Second, this study focused on individuals aged 60 years and older. Further research is needed to investigate the DoG-A in samples of young and middle-aged adults. Third, although we based our measure of DoG on four subscales, other domains are equally conceivable (e.g., health outcomes, social and spiritual experiences). Fourth, the Real money and Magazines subscales produced dichotomous variables. Although continuous variables are preferable, more real money and magazine trials would be required to construct a continuous measure, significantly increasing the costs of the study.

Conclusions and further research

Despite these limitations, this study represents an important step in research on delay of gratification in adulthood and older age. The DoG-A is a theoretically and empirically informed, behavioral and experimental measure of motivational self-regulation in adulthood. It has several advantages compared to existing measures of delay of gratification or delay discounting. First, the DoG-A includes

behavioral tasks and real rewards and not hypothetical choices (as in delay discounting procedures) or self-reports (as in DoG questionnaires). Second, it includes four reward types and not a single one as in previous DoG measures (Knolle-Veentjer et al. 2008; Wulfert et al. 2002). The present study as well as previous studies has shown the phenomenon of domain independence, which supports the need to use various reward types in order to achieve a broader and more realistic measurement of DoG. Third, it uses meaningful delay intervals (hours and weeks instead of minutes) and individualizes preferences for rewards (individually selected snacks and magazines). Forth, it is less affected than questionnaires by social desirability bias, because the goal of the test is not directly transparent. Rather, participants believe that the procedure is intended to measure their preferences and interests.

The question arises as to whether it might be enough to use the DoG-A Snacks subscale, which correlated more strongly with the self-regulation variables. As we have hypothesized above, the money and magazines may be less important to older Swiss people than snacks because of the relative wealth of this age group in Switzerland. However, these reward types might be more important in other samples. Therefore, it seems reasonable to use several reward types to capture self-regulatory ability in various samples. Future studies could also include somewhat different rewards, e.g., a higher amount of money could be used in the real money trial, or some sort of social experience as reward instead of the magazines.

Future research could also apply the DoG-A to clinical groups. As summarized above, self-control is impaired in several mental disorders. As a behavioral measure, the DoG-A captures a somewhat different aspect of self-control than do traditional self-report measures. It could thus prove beneficial for the assessment of patients by allowing a more comprehensive assessment. Moreover, it is known that self-regulatory abilities can be trained (Forstmeier and Rüddel 2007). Application of the DoG-A would make it possible to evaluate the effects of training programs on DoG and, in turn, on reduction of symptomatology. The delay discounting rate has been found to predict the treatment effect (Krishnan-Sarin et al. 2007; MacKillop and Kahler 2009); DoG-A could also be investigated as a predictor. The opportunities for applying the DoG-A in experimental studies are manifold. Finally, motivational ability measured on a behavioral level might be one of the components of successful aging (Forstmeier and Maercker 2008), which can be generally conceptualized as determined by the dynamics of gains and losses (Baltes 1997).

Acknowledgments This work was supported by a grant from the Hedwig Widmer Foundation, Zurich, Switzerland. We thank Sabine Maier, Stefanie Eicher, Katrin Blanke, Vesna-Maria Garstick, Kathleen Werner, Myriam Brandner, and Sandy Krammer for their assistance in recruitment and data collection.

References

Ainslie, G. (1975). Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin, 82*, 463–496.

Ayduk, O., Mendoza-Denton, R., Mischel, W., Downey, G., Peake, P. K., & Rodriguez, M. (2000). Regulating the interpersonal self: Strategic self-regulation for coping with rejection sensitivity. *Journal of Personality and Social Psychology, 79*, 776–792.

Baltes, P. B. (1997). On the incomplete architecture of human ontogeny: Selection, optimization, and compensation as foundation of development theory. *American Psychologist, 52*, 366–380.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioural change. *Psychological Review, 84*, 191–215.

Bandura, A. (1989). Self-regulation of motivation and action through internal standards and goal systems. In L. A. Pervin (Ed.), *Goal concepts in personality and social psychology* (pp. 19–85). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Baumeister, R. F., & Heatherton, T. F. (1996). Self-regulation failure: An overview. *Psychological Inquiry, 7*, 1–15.

Bembenutty, H. (1999). Sustaining motivation and academic goals: The role of academic delay of gratification. *Learning and Individual Differences, 11*, 233–257.

Bembenutty, H., & Karabenick, S. A. (1998). Academic delay of gratification. *Learning and Individual Differences, 10*, 329–346.

Bembenutty, H., & Karabenick, S. A. (2003). *Academic delay of gratification, future goals, and self-regulated learning*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL, April 21–25.

Chapman, G. B. (1996). Temporal discounting and utility for health and money. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 22*, 771–791.

Cohen, S., Kamarck, T., & Mermelstein, R. J. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior, 24*, 385–396.

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO personality inventory (NEO PI-R) and NEO five factor inventory. Professional Manual*. Odessa, FL: Psychological Assessment Resources.

Critchfield, T. S., & Kollins, S. H. (2001). Temporal discounting: Basic research and the analysis of socially important behavior. *Journal of Applied Behavior Analysis, 34*, 101–122.

de Wit, H. (2009). Impulsivity as a determinant and consequence of drug use: A review of underlying processes. *Addiction Biology, 14*, 22–31.

Derogatis, L. R. (1993). *Brief Symptom Inventory (BSI): Administration, scoring, and procedures manual* (3rd ed.). Minneapolis, MN: National Computers Services.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49*, 71–75.

Duckworth, A. L., & Seligman, M. E. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science, 16*, 939–944.

Feather, N. T. (1990). Bridging the gap between values and actions. Recent applications of the expectancy-value model. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of*

motivation and cognition: Foundations of social behavior (pp. 151–192). New York: Guilford Press.

Forstmeier, S., & Maercker, A. (2008). Motivational reserve: Lifetime motivational abilities influence cognitive and emotional health in old age. *Psychology and Aging, 23*, 886–899.

Forstmeier, S., & Maercker, A. (2010). *Belohnungsaufschub im Erwachsenenalter: Eine deutsche Version des Delay Discounting Tests von Kirby*. [Delay of gratification in adulthood: A German version of Kirby's Delay Discounting Test]. Manuscript submitted for publication.

Forstmeier, S., & Rüddel, H. (2007). Improving volitional competence is crucial for the efficacy of psychosomatic therapy: A controlled clinical trial. *Psychotherapy and Psychosomatics, 76*, 89–96.

Freund, A. M., & Baltes, P. B. (2002). Life-management strategies of selection, optimization, and compensation: Measurement by self-report and construct validity. *Journal of Personality and Social Psychology, 82*, 642–662.

Funder, D. C., & Block, J. (1989). The role of ego-control, ego-resiliency, and IQ in delay of gratification in adolescence. *Journal of Personality and Social Psychology, 57*, 1041–1050.

Funder, D. C., Block, J. H., & Block, J. (1983). Delay of gratification: Some longitudinal personality correlates. *Journal of Personality and Social Psychology, 44*, 1198–1213.

Green, L., Fry, A. F., & Myerson, J. (1994). Discounting of delayed rewards: A life-span comparison. *Psychological Science, 5*, 33–36.

Green, L., & Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin, 130*, 769–792.

Heerey, E. A., Robinson, B. M., McMahon, R. P., & Gold, J. M. (2007). Delay discounting in schizophrenia. *Cognitive Neuropsychiatry, 12*, 213–221.

Karoly, P. (1993). Mechanisms of self-regulation: A systems view. *Annual Review of Psychology, 44*, 23–52.

Kirby, K. N. (2009). One-year temporal stability of delay-discount rates. *Psychonomic Bulletin & Review, 16*, 457–462.

Kirby, K. N., Petry, N. M., & Bickel, W. K. (1999). Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. *Journal of Experimental Psychology: General, 128*, 78–87.

Knolle-Veentjer, S., Huth, V., Ferstl, R., Aldenhoff, J. B., & Hinze-Selch, H. (2008). Delay of gratification and executive performance in individuals with schizophrenia: Putative role for eating behavior and body weight regulation. *Journal of Psychiatric Research, 42*, 98–105.

Krishnan-Sarin, S., Reynolds, B., Duhig, A. M., Smith, A., Liss, T., McFetridge, A., et al. (2007). Behavioral impulsivity predicts treatment outcome in a smoking cessation program for adolescent smokers. *Drug and Alcohol Dependence, 88*, 79–82.

Krueger, R. F., Caspi, A., Moffitt, T. E., White, J., & Stouthamer-Loeber, M. (1996). Delay of gratification, psychopathology, and personality: Is low self-control specific to externalizing problems? *Journal of Personality, 61*, 107–129.

Kruglanski, A. W., Thompson, E. P., Higgins, E. T., Atash, M. N., Pierro, A., Shah, J. Y., et al. (2000). To "do the right thing" or to "just do it": Locomotion and assessment as distinct self-regulatory imperatives. *Journal of Personality and Social Psychology, 79*, 793–815.

Kuhl, J., & Fuhrmann, A. (1998). Decomposing self-regulation and self-control: The volitional components inventory. In J. Heckhausen & C. S. Dweck (Eds.), *Motivation and self-regulation across the life span* (pp. 15–49). Cambridge, UK: Cambridge University Press.

Levenson, H. (1974). Activism and powerful others: Distinctions within the concept of internal-external control. *Journal of Personality Assessment, 38*, 377–383.

MacKillop, J., & Kahler, C. W. (2009). Delayed reward discounting predicts treatment response for heavy drinkers receiving smoking cessation treatment. *Drug and Alcohol Dependence, 104*, 197–203.

Madden, G. J., Petry, N. M., Badger, G. J., & Bickel, W. K. (1997). Impulsive and self-control choices in opioid-dependent patients and non-drug-using control participants: Drug and monetary rewards. *Experimental and Clinical Psychopharmacology, 5*, 256–262.

Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. 5. The effect of delay and of intervening events on reinforcement value* (pp. 55–73). Hillsdale, NJ: Erlbaum.

Metcalfe, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: Dynamics of willpower. *Psychological Review, 106*, 3–19.

Mischel, W. (1961). Preference for delayed reinforcement and social responsibility. *Journal of Abnormal and Social Psychology, 62*, 1–7.

Mischel, W. (1974). Processes in delay of gratification. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 7, pp. 249–292). San Diego, CA: Academic Press.

Mischel, W., & Ebbesen, E. B. (1970). Attention in the delay of gratification. *Journal of Personality and Social Psychology, 16*, 329–337.

Mischel, W., & Metzner, R. (1962). Preference for delayed reward as a function of age, intelligence, and length of delay interval. *Journal of Abnormal and Social Psychology, 64*, 425–431.

Mischel, W., Shoda, Y., & Peake, P. K. (1988). The nature of adolescent competencies predicted by preschool delay of gratification. *Journal of Personality and Social Psychology, 54*, 687–699.

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science, 244*, 933–938.

Monterosso, J., & Ainslie, G. (1999). Beyond discounting: Possible experimental models of impulse control. *Psychopharmacology, 146*, 339–347.

Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., et al. (1989). The consortium to establish a registry for Alzheimer's disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology, 39*, 1159–1165.

Odum, A. L., & Rainaud, C. P. (2003). Discounting of delayed hypothetical money, alcohol, and food. *Behavioural Processes, 64*, 305–313.

Petry, N. M. (2002). Discounting of delayed rewards in substance abusers: Relationship to antisocial personality disorder. *Psychopharmacology, 162*, 425–432.

Rachlin, H. (1974). Self-control. *Behaviorism, 2*, 94–107.

Ray, J. J., & Najman, J. M. (1986). The generalizability of deferment of gratification. *Journal of Social Psychology, 126*, 117–119.

Reitan, R. (1958). Validity of the trail making test as an indication of organic brain damage. *Perception and Motor Skills, 8*, 271–276.

Reynolds, B. (2006). A review of delay-discounting research with humans: Relations to drug use and gambling. *Behavioural Pharmacology, 17*, 651–667.

Reynolds, B., de Wit, H., & Richards, J. B. (2002). Delay of gratification and delay discounting in rats. *Behavioural Processes, 59*, 157–168.

Richards, J. B., Zhang, L., Mitchell, S. H., & de Wit, H. (1999). Delay or probability discounting in a model of impulsive behavior:

Effect of alcohol. *Journal of the Experimental Analysis of Behavior, 71*, 121–143.

Rosenbaum, M., & Ben-Ari Smira, K. (1986). Cognitive and personality factors in the delay of gratification of hemodialysis patients. *Journal of Personality and Social Psychology, 51*, 357–364.

Rounds, J. S., Beck, J. G., & Grant, D. M. (2007). Is the delay discounting paradigm useful in understanding social anxiety? *Behaviour Research and Therapy, 45*, 729–735.

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A reevaluation of the life orientation test. *Journal of Personality and Social Psychology, 67*, 1063–1078.

Scheres, A., Lee, A., & Sumiya, M. (2008). Temporal reward discounting and ADHD: Task and symptom specific effects. *Journal of Neural Transmission, 115*, 221–226.

Schmidt, K.-H., & Metzler, P. (1992). *Wortschatztest (WST) [Vocabulary test (WST)]*. Weinheim: Beltz.

Scholz, U., Gutierrez Dona, B., Sud, S., & Schwarzer, R. (2002). Is general self-efficacy a universal construct? Psychometric findings from 25 countries. *European Journal of Psychological Assessment, 18*, 242–251.

Schultheiss, O. C., Yankova, D., Dirlikov, B., & Schad, D. J. (2009). Are implicit and explicit measures statistically independent? A fair and balanced test using the picture story exercise and a cue- and response-matched questionnaire measure. *Journal of Personality Assessment, 91*, 72–81.

Silverman, I. W. (2003). Gender differences in delay of gratification: A meta-analysis. *Sex Roles, 49*, 451–463.

Steinberg, L. (2007). Risk taking in adolescence: New perspectives from brain and behavioral science. *Current Directions in Psychological Science, 16*, 55–59.

Stroop, J. R. (1935). Studies of interference in serial verbal reaction. *Journal of Experimental Psychology, 18*, 643–662.

Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality, 72*, 271–322.

Ward, W. E., Perry, T. B., Woltz, J., & Doolin, E. (1989). Delay of gratification among Black college student leaders. *Journal of Black Psychology, 15*, 111–128.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063–1070.

Wechsler, D. (1997). *Wechsler adult intelligence scale (WAIS-III)* (3rd ed.). San Antonio: The Psychological Corporation.

Weller, R. E., Cook, E. W., Avsar, K. B., & Cox, J. E. (2008). Obese women show greater delay discounting than healthy-weight women. *Appetite, 51*, 563–569.

Witt, L. A. (1990). Delay of gratification and locus of control as predictors of organizational satisfaction and commitment: Sex differences. *Journal of General Psychology, 117*, 437–446.

Wulfert, E., Block, J. A., Santa Ana, E., Rodriguez, M. L., & Colsman, M. (2002). Delay of gratification: Impulsive choices and problem behaviors in early and late adolescence. *Journal of Personality, 70*, 533–552.

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., et al. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatry Research, 17*, 37–49.