DOI:

https://doi.org/10.1037/bar0000246

Save the Best for Last I: Young Adults Demonstrate Negative Time Preference- A Replication and Extension

Mariana I. Castillo, Shuyan Sun, Michelle A. Frank-Crawford, and John C. Borrero

University of Maryland, Baltimore County

Completion of this study and preparation of this manuscript were supported in part by grants from the UMBC Graduate Student Association.

The authors would like to thank Dr. Griffin Rooker, Dr. Jolene Sy, and the anonymous reviewers for their feedback on earlier drafts of this manuscript.

Direct correspondence to J. C. Borrero, Department of Psychology, UMBC, 1000 Hilltop Circle, Baltimore, Maryland 21250 (e-mail: jborrero@umbc.edu).

Mariana Castillo ORCID 0000-0002-7433-4324

Shuyan Sun ORCID 0000-0002-7512-2753

Michelle Frank-Crawford ORCID 0000-0003-4319-6317

John Borrero ORCID 0000-0003-0413-0892

Abstract

Generally, immediate outcomes are preferred to delayed outcomes, and in economics, this phenomenon is referred to as positive time preference. If positive time preference is normative, when asked to schedule a set of outcomes, a person should typically prefer a worsening sequence (i.e., choose to start with the best outcome, and end with the worst outcome). Several studies have shown that when a choice is among a *sequence* of outcomes, people typically exhibit negative time preference (i.e., they prefer an improving series of events). In the current studies college students responded to hypothetical questions via an online survey. Study 1a was a replication and extension of procedures described by Loewenstein and Prelec (1991, section II). Response patterns like those of Loewenstein and Prelec were observed, in that the percentage of participants who saved the best for last decreased when the interval between activities in the sequence increased. In Study 1b participants were surveyed about their preference for the order in which they would experience hypothetical outcomes with sequences of different sizes (e.g., three activities to sequence or eight). As array size increased, the percentage of participants who saved the best for last, or generated a perfectly improving sequence, decreased. In Study 2, 192 college students responded to questions involving categorically different outcomes (e.g., noxious stimuli, food, exercise, schoolwork, leisure) via an online survey. A smaller percentage of participants saved the best for last relative to prior studies, but the percentage was highest when sequences involved noxious stimuli or food.

Key words: preference for sequences, time, array, delay, save the best for last, negative time preference

Save the Best for Last I: Young Adults Demonstrate Negative Time Preference-

A Replication and Extension

Traditionally, studies of intertemporal choice (i.e., "decisions in which the moment of choice and the associated consequences are separated in time;" Lucci, 2013, p. 1) have involved the delay-discounting model. This model is concerned with how people evaluate or select single outcomes from an array of multiple possible outcomes, at a point in time (see Frederick et al., 2002, for a history on the discounted utility model). Although the rate of discounting varies within and across individuals, studies on delay discounting using considerably diverse procedures have suggested that as delays increase, the value of an outcome diminishes according to a hyperbolic function (Madden & Johnson, 2010). Thus, for intertemporal choice involving single-outcome prospects (i.e., only one outcome is selected from the array), delayed outcomes are generally valued less than those that are not delayed (across species, and within humans of different characteristics; Ainslie, 1975). In economics, this phenomenon is referred to as *positive time preference* (Olson & Bailey, 1981).

If positive time preference is evinced in all of one's choices, then that would imply that when asked to schedule a set of outcomes, a person should typically prefer a worsening sequence (i.e., choose to start with the best outcome, followed by the second best, and ending with the worst outcome). Several studies have specifically focused on this problem and have shown that when a choice is viewed as part of a sequence of outcomes, people typically exhibit *negative time preference* (i.e., prefer an improving series of events; Loewenstein & Prelec, 1991; Magen et al., 2008). *Negative time preference* refers to a general preference for improvement, which can manifest in a perfectly improving sequence or in a sequence that simply ends with the most preferred option.

Taken together, the existing research shows that the framing of the choice affects how people behave (Tversky & Kahneman, 1981). More specifically, when the frame draws attention to individual components of the choice (e.g., would you like to read *or* watch a television show now?), people tend to exhibit positive time preference (Frederick et al., 2002), choosing their most preferred immediately; when the decision frame draws attention to the sequential aspect of the choice (e.g., which would you like to do *first*: read or watch a television show?), people tend to exhibit negative time preference, saving their most preferred for last, in a sense "delaying gratification" (Loewenstein & Prelec, 1991; Magen et al., 2008).

With a brief three-question survey, Loewenstein and Prelec (1991, 1993) investigated how manipulating a specific parameter of a sequence, namely the timing of the outcomes, influenced people's choices in a predictable manner (i.e., showing positive or negative time preference). In the first question, the series of outcomes unfold over a fairly short period (2 weeks), with a brief (1 week) interval between the possible outcomes. In this case, 90% of subjects reported negative time preference, choosing to save the best for last. In the second question, there was a longer interval between both outcomes (i.e., the first outcome would be experienced immediately, whereas the second one would be experienced after a 6-month delay). Therefore, participants could potentially view this choice as two independent single-outcome choices instead of a sequence. Indeed, a significantly smaller percentage of participants (52%) demonstrated negative time preference given the long absolute delay. In the third question, both outcomes were delayed about 6 months, but the sequence interval (i.e., the timing between the outcomes) was once again reduced to 1 week. Like in the first question, most participants (83%) exhibited negative time preference. These results suggested that when outcomes were presented as a sequence, negative time preference was most likely, but the percentage of participants who preferred an improving sequence decreased when the interval between outcomes was increased.

Studies involving different methods (e.g., differences in the number of outcomes in the sequence, whether outcomes were hypothetical or real, or whether participants' ratings involve prospective evaluations before they experience the sequence or retrospective evaluations of sequences they have already experienced), have documented preference for improving sequences across a wide variety of outcomes including money (e.g., Duffy et al., 2015; Magen et al., 2008), pain (e.g., Ariely & Carmon, 2000), annoying sounds (e.g., Ariely & Zauberman, 2000), restaurant meals (e.g., Loewenstein & Prelec, 1991, 1993), health outcomes (e.g., Chapman, 1996, 2000), affective stimuli (e.g., Drolet et al., 2011), and news (e.g., Legg & Sweeny, 2014; Marshall & Kidd, 1981).

In addition to showing that the timing or spacing of outcomes influenced people's choices predictably, prior research has also suggested that if presented with more than one outcome of the same valence (i.e., perceived value), people preferred to spread such outcomes over time rather than concentrating them (Loewenstein & Prelec, 1993). For example, when deciding the order of four dinners, two at home and two at a fancy restaurant, people preferred to intersperse dinners at home with dinners at a fancy restaurant, rather than accumulating the dinners at the fancy restaurant in one extreme of the sequence. Because Loewenstein and Prelec (1993) compared five-outcome sequences with only three distinct valences (i.e., two low preferred, two moderately preferred, and one highly preferred) whether sequences involving outcomes of distinct valence would result in similar preferences for overall improvement and spreading is a matter not yet resolved (Chapman, 1996; Guyse et al., 2002). Further, it is unclear whether distinct preference patterns would emerge with sequences of different array sizes (i.e., different total number of outcomes).

Evaluations of meaningful parameters, such as sequence size or category of outcomes, are needed to determine the prevalence of “saving the best for last” and preference for improving sequences. Although not explicitly compared yet, the number of outcomes in a sequence may influence people’s choices for the temporal ordering of the sequence. Studies on preference for sequences of affective events have found preference for improving sequences when they involved only three outcomes (Drolet et al., 2011), but preference for mixed sequences when sequences involved 30 outcomes (Löckenhoff et al., 2012). Conceivably, preference for improvement is most probable given smaller sequences, and mixing preferred with less preferred outcomes is more probable for larger sequences. This conclusion must be stated speculatively, however, because the methods in the two cited studies differed in more ways than just the number of outcomes in the sequences.

Furthermore, some studies explicitly evaluating methodological variations have found that the preference for improvement may not be as robust as suggested by the extant literature on preferences for sequences (Frederick & Loewenstein, 2008). For example, Frederick and Loewenstein (2008) found that when the task involved creating sequences by allocating a fixed quantity of goods over a series of time periods (e.g., schedule 20 events over a 5-year period), respondents allocated outcomes evenly about half of the time. The rest of the time, improving or worsening sequences were equally likely. In a notable study, Andrade and Hackenberg (2012) conducted a cross-species analysis of choices between reinforcer sequences to discern some of the reasons for the often-reported discrepancy in the results of choice experiments with human and nonhuman subjects. Andrade and Hackenberg posited that the discrepancy in choice may in fact stem from methodological differences. An important procedural difference is the format of presenting choice (e.g., single shot versus repeated exposure, hypothetical versus real contingent outcomes). Therefore, in Experiment 2, human participants made repeated pairwise choices among three types of sequences types (i.e., standard, improving, or worsening), each of which provided the same overall reinforcement rate (i.e., four tokens), but differed in the patterning of the intertemporal delays for the token presentation (i.e., equal delays, successively decreasing delays, or successively increasing delays). The results were generally consistent with models of temporal discounting, with participants preferring the worsening sequences (i.e., those that started with the shortest delay). The same participants then answered a questionnaire like that of Loewenstein and Prelec (1993). Results of the questionnaire were consistent with prior questionnaire studies but did not show any clear relation to their choice patterns in Experiment 2. As Andrade and Hackenberg suggested, “different methods may recruit different decision-making strategies” (pp. 62-63). There are several possible differences that could have exerted control over the participants’ behavior. These include differences in the type of outcome (e.g., real versus hypothetical), the number of response opportunities (e.g., repeated exposure versus single shot), the manner in which the value of the sequence is manipulated (e.g., delay differences versus magnitude or qualitative differences) and the type of outcomes (e.g., brief video clips versus hypothetical meals or vacations).

Studies have found that the effect of delays on time preference may depend on the category of the outcome—the type of outcome. Kumar et al. (2014) found that participants derived pleasure from anticipating experiences (e.g., concerts, meals out, vacations), and waiting for experiences tended to be more enjoyable than waiting for material goods (e.g., clothing, gadgets). Although there is evidence that individuals sometimes prefer to delay a desirable outcome (Loewenstein, 1987; Kumar & Gilovich, 2016), preference for reducing the delay to aversive outcomes seems more widespread (Cook & Barnes, 1964; Story et al., 2013). When given the option to choose between an immediate or delayed inevitable shock, the majority (approximately 78%) of participants opted to receive more-painful shocks right away rather than wait for less painful shocks following a brief delay (Story et al., 2013). Story et al. posited that the anticipation of forthcoming and inevitable pain is so aversive in and of itself that people will pay a significant price, in the form of more physical pain, to avoid it.

Moreover, a study evaluating preferences for temporal sequences of categorically-different real experiences, found that participants seemed to prefer improving sequences for aversive physical experiences (i.e., decreasing intensities of shock) but mixed sequences for physical effort and monetary payouts (Löckenhoff et al., 2019). Taken together, these results suggested that negative time preference in sequences is not universal, and that time preference may be influenced by the framing of the choice and by the category of outcomes involved.

Thus far, the focus of this introduction has revolved around contextual factors that have an impact on choice. Nevertheless, individual characteristics may also influence choice and time preference as it relates to sequences. Among such characteristics, one is relevant to the current investigation, and has received limited attention as it relates to preference for sequences: impulsiveness.

Although time preference and impulsiveness are different constructs, we hypothesize that they may be related. Delay discounting, an example of positive time preference, is considered a measure of impulsivity (Odum, 2011), and discounting has been positively associated with self-report measures of impulsiveness (de Wit et al., 2007). Therefore, it is possible that negative time preference related to the sequence frame is also associated with impulsiveness and self-control. Presumably, highly impulsive individuals may be less likely to exhibit negative time preference, or less likely to save the best for last, relative to less impulsive individuals.

The current manuscript describes two studies designed as a replication and extension of Loewenstein and Prelec's (1993) seminal paper on preference for sequences. Specifically, Study 1 represents a two-part study investigating whether the timing of choices and the number of outcomes in a sequence, influence participants' choice to save the best for last. Study 2 sought to expand on existing research that has evaluated categorical features of outcomes in a sequence and their effect on time preference. Understanding some of the contextual factors that influence human decision making is squarely within the domain of behavior analytic research and practice, the translational studies outlined herein represent some preliminary attempts to better understand a subset of these potential factors.

Study 1

The purpose of Study 1, a two-component study, was to replicate and extend the research by Loewenstein and Prelec (1993), investigating whether the timing of choices and the number of outcomes in a sequence, influence participants' choice to save the best for last. In Study 1a, we replicated and extended procedures described by Loewenstein and Prelec (1991, section II) with college students responding to questions about hypothetical events on an online survey. Next, in Study 1b we surveyed college students about their preference for the order in which they would experience hypothetical outcomes with sequences of different sizes (e.g., three activities to sequence or eight).

Method

Participants and Setting

A total of 279 undergraduate college students participated in this study (see Table 1 for demographic information). By comparison, little is known about participants in Loewenstein and Prelec's 1991 and 1993 studies. Readers learned that participants were "48 visitors to the Museum of Science and Industry in Chicago." For the present study, participants were recruited from psychology courses and invited to participate in the study via a psychology participant pool, for extra credit. The study was completed over a nine-month period, from March through December. This research project was reviewed and approved by the UMBC Institutional Review Board.

Measures and Procedures

After signing up for the study via the psychology department's research participation website, participants were provided a link that they used to access and respond to an online survey in Qualtrics®XM (Qualtrics, Provo, UT). The survey took approximately 21 min to complete, on average. Following completion of the survey, extra credit was automatically awarded on the psychology department's research participation website.

At the beginning of the survey, participants were provided with the following statements: "In this survey, we will ask you a number of questions about how you prefer to plan or schedule different activities. For all questions, read the description carefully, and please ignore any scheduling considerations. That is, assume you don't have any prior commitments and you can actually do all the activities in the proposed time period."

Study 1a. Time between events and sequence duration as predictors of positive time preference. The first three questions of Study 1a involved a *systematic*, rather than a *direct*, replication (Sidman, 1960) of procedures described by Loewenstein and Prelec (1991, 1993). Although the first three questions were identical to those of Loewenstein and Prelec, we used a different method (i.e., Qualtrics®XM) to survey participants. Moreover, participant recruitment differed in both studies. Specifically, we recruited undergraduate students from a university in the mid-Atlantic region and Loewenstein and Prelec recruited visitors to the Museum of Science and Industry in Chicago. The first page of the survey consisted of the same three questions used by Loewenstein and Prelec. The scenario and questions appeared on the screen at the same time and are described next.

> Imagine you must schedule two weekend outings to a city where you once lived. You do not plan on visiting the city after these two outings.
>
> You must spend one of these weekends with an irritating, abrasive aunt[1] who is a horrendous cook. The other weekend will be spent visiting former work associates whom you like a lot. From the following pairs, please indicate your preference by checking the appropriate line.
>
> (1) Suppose one outing will take place this coming weekend, the other the weekend after.
> - This weekend **friends**; next weekend **abrasive aunt**
> - This weekend **abrasive aunt**; next weekend **friends**
>
> (2) Suppose one outing will take place this coming weekend, the other in 6 months (26 weeks).
> - This weekend **friends**; 26 weeks from now **abrasive aunt**
> - This weekend **abrasive aunt**; 26 weeks from now **friends**
>
> (3) Suppose one outing will take place in 6 months (26 weeks from now), the other the weekend after (27 weeks from now).
> - 26 weeks from now **friends**; 27 weeks from now **abrasive aunt**

[1] Because the term "abrasive aunt" was used in previously published works, we used the same term so as to avoid further changes in a replication study. However, this term should be updated in future studies to reduce any misogynistic overtones.

- 26 weeks from now **abrasive aunt**; 27 weeks from now **friends**

The second page of the survey described next, was an extension of Loewenstein and Prelec (1991, 1993) and included three similar questions modified slightly to reflect smaller time windows (i.e., within the same hour, same day, or a day apart).

> Imagine you are visiting a city where you once lived. You do not plan on visiting the city after this trip. You must schedule two outings during your trip. You must spend one of these outings with an irritating, abrasive aunt who is a horrendous cook. The other outing will be spent visiting former work associates whom you like a lot. From the following pairs, please indicate your preference by checking the appropriate line.
>
> (4) Suppose one outing will take place from 10:00 to 10:25, the other 10:30 to 10:55.
> - 10:00 – 10:25 **friends**; 10:30 – 10:55 **abrasive aunt**
> - 10:00 – 10:25 **abrasive aunt**; 10:30 – 10:55 **friends**
>
> (5) Suppose one outing will take place at 9:00 am, the other at 1:00 pm on the same day.
> - 9:00am **friends**; 1:00pm **abrasive aunt**
> - 9:00am **abrasive aunt**; 1:00pm **friends**
>
> (6) Suppose one outing will take place on the 21st of the month, the other on the 22nd of the same month.
> - 21st of the month **friends**; 22nd of the month **abrasive aunt**
> - 21st of the month **abrasive aunt**; 22nd of the month **friends**

Study 1b. Number of items in the sequence as a predictor of positive time preference. The second part of the survey was a variation of the survey used in Study 1 of Loewenstein and Prelec (1993) and consisted of questions in which participants selected the order in which to experience the hypothetical outcomes in sequences of different sizes. Participants were first provided with this instruction:

> For the second part of the survey, you will need to think of brief activities that you like to different degrees and can be finished within a day.

> Think of something you really enjoy doing with people you really like, a favorite activity.
>
> Think of things you enjoy, but do not love, doing.
>
> Think of something you do not enjoy doing, a boring and not at all fun activity.
>
> In the space below, list 8 activities that you like to different degrees. Rank them from most preferred (1) to least preferred (8).

Participants were required to answer the question entirely (i.e., type something in each of the eight text boxes), before proceeding with the survey. The activities listed by the participant were presented as a reminder throughout the rest of the survey. For the first 153 participants, the sequence questions were presented in an increasing order (starting with two outcomes and ending with eight). In acknowledgment of order effects (Chapman, 2000), the second 126 participants were randomly assigned to one of three conditions in which the array sizes were presented in three possible orders: an increasing order (defined above), a decreasing order (starting with eight and ending with two), or one randomly generated order (experiencing the following order of array sizes: 7, 4, 8, 3, 5, 6, 2). Each sequence consisted of outcomes of similar duration that could all be experienced within the same day without regard to scheduling issues (i.e., the participant could experience the outcome whenever chosen, and should therefore ignore scheduling considerations).

As in Study 1 of Loewenstein and Prelec (1993), the arrays consisted of boxes of different heights, which represented the participant's preference for different leisure activities. A sample diagram with a randomly generated order was provided, and the participants selected the order in which they would like to experience each activity. An example of the instructions for a three-item array is presented below:

> Imagine it is early morning and you are contemplating how to schedule your day. Your situation is depicted in the diagram below (a diagram was displayed for participants as well as the accompanying text below).
>
> The height of the box represents how much you enjoy the activity. The taller the box, the more you like that activity. In the diagram, you are doing the most pleasurable thing second.
>
> The medium height box represents a moderately pleasurable activity. Think of something you enjoy, but do not love, doing. In the diagram, you are doing the moderately pleasurable activity first.
>
> The flat line represents a boring and not at all fun activity. Think of something you do not enjoy doing. In the diagram, you are experiencing the most boring activity last.
>
> As a reference, here is the list of activities you previously generated, in your preferred order: [the eight activities that had been provided at the beginning of Study 1b, would appear here].
>
> Imagine it is early in the morning, how would you like to schedule your day? Select the order in which you would like to experience each activity (first, second, last).

The answer section consisted of a matrix, in which the different boxes were presented as column headers, and the order options (i.e., first, second, …last) as line headers. The participants selected the radio buttons within the matrix that corresponded to their choice.

Results

Study 1a

How does the timing of sequences influence saving the best for last? Table 2 includes the data on the direct replication and systematic extension of Loewenstein and Prelec (1991, 1993), presented as the percentage of participants who selected to see their abrasive aunt first, and their friends last for all six questions in Study 1a. For almost all the time windows, the majority (range, 68% to 86%) chose to schedule the outing with their friends last in the sequence.

The only exception was the second question, in which the interval between both outings was much greater (6 months), and only 37% of participants exhibited negative time preference

The patterns observed in our results in the first three questions are like those reported by Loewenstein and Prelec (1991, 1993), in that the highest percentage of participants who showed negative time preference was observed for question 1, the second highest was observed for question 3, and the smallest was observed for question 2. The researchers conducted a chi-squared goodness of fit test to compare the percentage of participants who showed negative time preference in each of the first three scenarios to those same percentages reported by Loewenstein and Prelec. A similar percentage of participants in both samples displayed negative time preference in question 1, when the outing options were this weekend or next weekend (i.e., differences in the percentages were not statistically significant: χ^2 (1) = 2.704, p = .1). In the present sample, however, a significantly smaller percentage of participants chose to save the best for last in question 2, when the outings were spaced six months apart (χ^2 (1) = 3.854, p = .049), and in question 3, when the interval was again reduced to one week, but both outings were delayed by six months (χ^2 (1) = 4.391, p = .036).

Although the patterns were similar across the two studies, significantly fewer participants in our study chose to save the best for last when delays were involved (i.e., second and third questions), as compared to that of Loewenstein and Prelec (1991, 1993). For the last three questions in Study 1a, which involved smaller time windows, a large percentage of participants, chose to save the best for last, just like they did for the first question.

Study 1b

In Study 1b, participants experienced the change in array size in one of three different orders: increasing order (starting with two outcomes and ending with eight), decreasing order (starting with eight and ending with two), and one randomly generated order (experiencing the following order of array sizes: 7, 4, 8, 3, 5, 6, 2). Results of Study 1b are presented next. When the differences across groups were not statistically significant, the results were combined into a single group. When differences were observed across groups, results were described per group.

How does array size influence saving the best for last? Figure 1 shows the percentage of participants who saved the best for last (i.e., who chose to put the tallest box last in the sequence) for each array size. Combining all array sizes, more than half of participants M = 62.88%, range, 56.99% to 67.74%) of participants chose to save the best for last. The experimenters conducted a linear regression analysis to determine how well array size could predict the percentage of participants who saved the best for last. Results of the regression analysis were significant (F (1, 5) = 19.08, p = .007), indicating that there is a tendency for the percentage of participants who saved the best for last to change as array size changes. The regression equation $\hat{Y}$ = -1.51X. + 70.43 resulted in a R^2 of .79. That is, 79% of the variance in percentage of participants who saved the best for last was predicted from array size. The negative slope of the linear regression (B = -1.51, p = .007, 95% CI [-2.39, -.62]) indicates that as array size increased, the percentage of participants who saved the best for last decreased.

How does array size influence preference for a perfectly improving sequence? Does the order of exposure matter? To avoid redundant data, results for the array size of two are not included in these analyses, because saving the best for last is equivalent to generating a perfectly improving sequence when there are only two outcomes in the sequence. Table 3 shows the percentage of participants who generated a perfectly improving sequence in each group, for array sizes 3 through 8. For each array size, we conducted chi-squared goodness of fit tests to compare the percentage of participants who generated perfectly improving sequences in each group, and results are depicted in the last columns of Table 3. The percentage of participants who generated perfectly improving sequences was not statistically different across order groups for array sizes 3 through 5. Chi-squared tests for arrays of six, seven, or eight outcomes suggested that the difference in percentages across groups was statistically significant ($p < .05$). Therefore, these data suggest that the effect of array sizes varied by the order of exposure. More specifically, the percentage of participants who generated a perfectly improving sequence for array sizes 6, 7, and 8, was greatest when those were presented toward the end of the survey. In other words, individuals were more likely to generate a perfectly improving sequence in the larger array sizes when those larger arrays were experienced at the end of the survey, as was the case for the increasing group.

For each experience order group, the experimenters conducted linear regression analyses with percentage of participants who generated a perfectly improving sequence as the dependent variable and array size as the predictor. For this analysis, we also excluded data for the array of two outcomes. Results of the regression analysis indicated that the effect of array size on the percentage of participants who generated a perfectly improving sequence was significant for the groups that experienced the array sizes in increasing ($F(1, 4) = 15.74$, $p = .017$) and decreasing order ($F(1, 4) = 48.37$, $p = .002$), but was not significant for the group that experienced the random order ($F(1, 4) = 7.07$, $p = .056$). The regression equations for predicting the percentage of participants who generated a perfectly improving sequence using array size, for each group are depicted in Table 3. The R^2 for these equations for the increasing, decreasing, and random group indicate that 79.7% of the variance in percentage of participants who generated a perfectly improving sequence was predicted from array size for the increasing group, 92.4% for the decreasing group, and only 63.9% for the random group. The negative slope of the linear regressions for the increasing and decreasing groups indicate that as array size increased, the percentage of participants who generated a perfectly improving sequence decreased.

Table 4 summarizes the percentage of all participants who saved the best for last, as well as the two most common sequence patterns generated for each array size, namely perfectly improving and perfectly worsening sequences. In all array sizes, the most commonly generated sequence was a perfectly improving sequence, indicative of negative time preference. The second most common sequence was a perfectly worsening sequence, indicative of positive time preference (i.e., delay discounting). For the array size of 3, the third most commonly generated sequence (16.49%) involved scheduling the moderately preferred outcome first, followed by the least preferred, and ending in the most preferred. All other sequences, across all array sizes, were generated by fewer than 9% of the sample, therefore, they will not be described specifically, but can be made available upon request.

Looking at array sizes 3 – 8, for the entire sample combined, a mean of 37.63% (range, 31.54% to 48.03%) of participants generated a perfectly improving sequence, demonstrating negative time preference. A smaller percentage of participants generated a perfectly improving sequence for all array sizes compared to the percentage of participants who saved the best for last, regardless of the order in which it was experienced. With the percentages ranging from 31.54% for the sequence of eight outcomes, to 48.03% for the sequence of three outcomes, both percentages were lower than the smallest percentage who saved the best for last (56.99% for the sequence of seven outcomes). Chi-squared tests comparing the percentages for each array size indicated that the differences were all statistically significant ($p < .001$). Taken together, approximately 62% of participants arranged selections to save the highest ranked item for last, whereas a smaller subset of those individuals also arranged all selections to progress from “worst to best” (37.63%).

Study 2

The purpose of Study 2 was to investigate the likelihood of finding negative time preference, and saving the best for last, in sequences of categorically-different hypothetical outcomes and their relation to self-report measures of impulsiveness. Specifically, we aimed to determine whether college students’ likelihood of saving the best for last differed based on the category of the outcomes in the sequence (categories are explicitly described below). We also aimed to determine whether impulsiveness, and specifically non-planning impulsiveness (i.e., a lack of “future thinking”), is related to time preference in sequences. A thoroughgoing understanding of the roles that outcome categories have on time preference could have implications for a wide range of potential applications: (a) promoting pro-health behavior (e.g., exercise), (b) facilitating tolerance or compliance with painful medical procedures, (c) enhancing employee performance in organizational settings, (d) improving educational outcomes, and (e) identifying preferred items (and potential reinforcers) for behavioral interventions for persons with intellectual and developmental disabilities.

Method

Participants and Setting

A total of 192 undergraduates participated in Study 2 (see Table 5 for demographic information). Participants were recruited from psychology courses and invited to complete an online survey via a psychology participant pool, for extra credit. Recruitment and data collection were completed over a two-month period, from mid-March through mid-May.

Measures and Procedures

Overview. Participants who contacted the research team were asked to respond to an online survey in Qualtrics®XM (Qualtrics, Provo, UT). The survey consisted of two parts completed consecutively in a single sitting: the Sequence Survey and the Barratt Impulsiveness Scale (BIS; Patton et al., 1995). The entire survey took approximately 21 min to complete, on average. The impulsiveness scale was included to correlate with the results of the Sequence Survey and better understand the relation between time preference and impulsivity, as highly impulsive individuals may be less likely to exhibit negative time preference, or less likely to save the best for last, relative to less impulsive individuals.

Part 1: Sequence Survey. The survey was similar to that used in Study 1b, except that during Study 2 verbal descriptions were used to represent the stimuli, rather than abstract graphical representations. The survey for the current study consisted of six sequences, each involving categorically-different outcomes (see **Appendix**

for a text copy of the survey questions). Categories included: (a) noxious outcomes, (b) food, (c) leisure, (d) schoolwork, (e) exercise, and (f) a mixed category. Following informal interviews with undergraduate students, the study team members collectively generated a list for each category that included items with which most undergraduate students have presumably had experience. The food list included strawberries, broccoli, tomatoes, chips, and crackers. The leisure list included reading a magazine, watching YouTube®, spending time with their phones, playing a paper game (e.g., Sudoku or crossword puzzles), and playing a game on a computer or tablet. The schoolwork list included three- and four-digit multiplication problems, long division, transcribing hand-written notes with a computer, reading comprehension, and building a model with plastic rods and connectors. The exercise list included jumping jacks, sit-ups, push-ups, running in place, and squats. Unlike the other categories, which included items with which most students have likely had experience, the list of noxious experiences was based on stimuli that are often used to study pain and sensory perception in clinical and laboratory settings (Yarnitsky & Pud, 2004). The items selected for the noxious category have been used extensively for neurological sensory testing with humans, including children with intellectual and developmental disabilities (Barney et al., 2017; Mitchell et al., 2004; von Baeyer et al., 2005). Specifically, noxious stimuli included submerging the hand in very cold water (approximately 1^{o} C, or ~ 33^{o} F), having the skin on the inner wrist pricked with a small plastic pin, experiencing the equivalent of approximately 1.8 kg of pressure on their inner wrist, having the end of a thick and stiff nylon thread pressed repeatedly against the skin on the inner wrist (30 times in 30 s), and having a warm probe (50^{o} C, or 122^{o} F) touch the skin of the inner wrist. The mixed category included the top-ranked item from each of the other categories.

Participants were first provided with a list of items in each category and asked to rank them in order of preference. After ranking the items in each categorical list (e.g., noxious outcomes 1-5), participants were asked to select the order in which they would like to experience the five different hypothetical outcomes in a sequential order. Category presentation was randomized across participants (some participants experienced the noxious category first, others the food category first, and so on), but the mixed category was always presented last.

An array of five outcomes was selected for individuals to sequence in a preferred order based on the results of Study 1. In that study, the sequence of five items was the largest array size in which the percentage of participants who generated a perfectly improving sequence did not differ based on the order in which it was presented in the survey (χ^2 (2) = 3.970, p = .137). Furthermore, when combining across experience order groups, the array size of five was significantly different from that of the sequence of three items (χ^2 (1) = 6.5779, p = .01).

Part 2: Barratt Impulsiveness Scale. Immediately following completion of the Sequence Survey, participants were asked to complete the BIS-11 (Patton et al., 1995) via Qualtrics. The BIS-11 is a 30 item self-report measure designed to assess impulsiveness (Patton et al., 1995). The BIS was originally published 60 years ago (Barratt, 1959), and at the time of its 50th anniversary, it was reportedly the most commonly administered self-report instrument for the assessment of impulsiveness in research and clinical settings (Stanford et al., 2009). Principal component analysis of the BIS-11 produced six first-order factors, and three second-order factors. These second-order subscales include Motor Impulsiveness (first-order factors motor and perseverance), Non-Planning Impulsiveness (first-order factors self-control and cognitive complexity), and Attentional Impulsiveness (first-order factors attention and cognitive instability). Self-report measures of self-control have been found to have high convergent validity (Duckworth & Kern, 2011), and scores on the non-planning subscale of the BIS-11 have been associated with delay discounting in previous studies (de Wit et al., 2007). In the current study, we conducted analyses using the overall score of the BIS-11, as well as the score in the Non-Planning Impulsiveness scale. Cronbach's alpha of the BIS-11 in this study was reported in Table 7.

Results

How do the categorical features of items in the array influence saving the best for last? To answer this question, we determined whether a participant selected the option they previously ranked #1 to be last in the sequence. Sequences consisted of five items. The probability of choosing the item ranked #1 in the last position, if participants were randomly clicking on the response options, was 0.2. Conversely, the probability of selecting one of the other four options last in the sequence was 0.8. We conducted a chi-squared test to determine whether the percentage of participants who saved the best for last in any given category was different from what would be expected had participants answered randomly. The results of this analysis are depicted in Table 6. In all categories, less than a third of the participants (range, 16.67% to 30.73%) chose to save the best for last in the sequence. The chi-squared test suggests that in only two categories, namely food and noxious stimuli, did the percentage of participants differ from what would be expected if participants had responded randomly. Specifically, when the sequences involved food or noxious stimuli, participants were more likely to save the best for last.

How does impulsiveness correlate with the preference for sequences? For each sequence the researchers calculated a Spearman rank order correlation between the relative value of the outcome, determined by the rank assigned by the participant and the order in which it was selected. This correlation captures individual variations in sequence preferences and is appropriate for rank data. The resulting scores ranged from −1 (perfectly improving: rank is inversely related to position in sequence) to 1 (perfectly worsening: rank is directly related to position in sequence) with scores closer to zero indicating a preference for interspersing higher and lower preferred outcomes. A Spearman correlation coefficient was obtained for each category. Two Pearson correlations were calculated. The first correlation was between each participant's overall score in the BIS-11 and the coefficient of the Spearman rank-order correlation between preference rank and position order in the sequence. The second correlation was between each participant's score in the non-planning subscale of the BIS-11 and the Spearman rank-order correlation between preference rank and position order in the sequence. Because the overall score and the non-planning subscale of the BIS-11 as well as Spearman correlations are continuous variables, Pearson correlation is appropriate to estimate the strength of relations between them. **Table 7** depicts Cronbach's alpha reliability of the BIS-11 and the subscales, as well as the mean score and standard deviation for each. **Table 8** depicts the Pearson's correlation coefficients between the Spearman rank-order correlation and the score in the entire BIS-11 or just the non-planning subscale. No significant associations were observed.

We conducted further exploratory analyses by separating the results into four different groups: (a) participants who generated diverse sequences but saved the best option to the end, (b) participants who generated perfectly improving sequences (i.e., worst-to-best), (c) participants who generated perfectly worsening sequences (i.e., best to worst), and (d) all other participants. Impulsiveness scores did not differ significantly across groups. These data are available upon request.

Discussion

The current set of studies aimed to replicate and extend the research by Loewenstein and Prelec (1991, 1993), with the addition of a study that incorporated individually identified stimuli into the survey, as well as a comparison of stimulus categories. Study 1 aimed to investigate whether the timing of choices and the number of outcomes in a sequence influence participants' choice to save the best for last. In Study 1a, we presented college students with the same questions that Loewenstein and Prelec (1991, 1993) used and a variation of those questions. In the current study, observed response patterns were like those observed by Loewenstein and Prelec and by Andrade and Hackenberg (2012), when using questionnaires. Specifically, the percentage of participants who exhibited negative time preference was highest when both options occurred relatively soon and with a one-week interval between options, the second highest percentage was observed when the interval remained the same but both options were delayed by six months, and the lowest percentage was observed when the interval between options was increased to six months. There was, however, a significant reduction in the percentage of participants in the current sample who exhibited negative time preference when there were delays involved, compared to those reported by the other two studies. Several potential reasons can be considered to explain why a smaller percentage of the current sample chose to save the best for last when there was a delay involved. For example, differences in sample demographic characteristics may account for the difference in responding. Perhaps there is something inherently different between undergraduate students in a university in the mid-Atlantic region and visitors to the Museum of Science and Industry in Chicago. It is possible that it is a cohort effect. Contemporary wisdom (Gleick, 2000) seems to suggest that with technological advances aspects of our society have, in a sense, sped up. Along with the internet, which became popularized in the early-to-mid 1990s with the invention of the world-wide web (Leiner et al., 2009) came substantive advances that have reduced the delay to many things like information, music, movies, mail, deliveries of online shopping, and communication, we have therefore become more sensitive to delays or more impatient (American Psychological Association, 2008). This sensitivity to delays in sequence outcomes is an area suitable for further direct and systematic replication.

Three additional questions were included in Study 1a, referencing additional time frames below that of the original questions from Loewenstein and Prelec (1991). The new and adjusted sequences included timeframes on the order of min, hours, and days. If the time frame truly affects sequence, then there should probably have been some sort of parametric difference between these conditions. It seems that more people should save the best for last when there are min apart and there should be progressively less people as the delay increases. Nevertheless, there seemed to be some sort of ceiling effect, which may represent an area for future research. Perhaps there is a different preparation that would allow for a parametric analysis of timeframes.

Interestingly, the percentage of the current sample who saved the best for last in Study 1b, when there were only two outcomes in the array, was smaller than it was in Study 1a (on all questions except number 2), which also involved sequences of only two outcomes. Methodological differences may account for these discrepancies. In Study 1a, participants were provided a clear description of the different outings in the sequence. In Study 1b, however, a list was not provided, and participants were prompted to think of brief activities that they like to varying degrees including favorite activities, things they enjoy but do not love doing, and boring activities they do not enjoy. Therefore, in Study 1b, there was variability, within and across participants, in the type of outcomes that they ranked during the survey. For example, some participants included only items that could be considered leisure activities such as listening to music, spending time with friends, going on nature walks, drawing, and bowling. Whereas other participants included activities such as taking out the trash, doing chores, watching a bad movie, doing homework, going to the dentist, or going to a funeral. Perhaps there are qualitative aspects of the outcomes imagined that exerted control over their responding. Furthermore, although there is precedent in the literature for using bars of different heights to represent relative quantitative or qualitative differences in options from which participants choose (Chapman, 2000; Löckenhoff et al., 2019; Loewenstein & Prelec, 1993), it is possible that the boxes in the diagram may have been too abstract for some participants to visualize the hypothetical activities they were to schedule. A more circumscribed list of outcomes provided to participants to rank in Study 1b may have produced very different results.

In Study 2, we investigated how young adults choose to schedule outcomes of different categories, and approximately a quarter of participants chose to save the best for last. This is quite different from the results of Study 1b in which 64.87% of participants saved the best for last when arranging a sequence of five hypothetical outcomes. One possible explanation for this discrepancy is the procedural variations in the studies. Specifically, in Study 2, participants ranked and scheduled textual descriptions of different items, activities, or stimuli. Conversely, in Study 1b, participants were not provided with specific items, activities, or stimuli to rank, but were prompted to think of brief activities they liked to different degrees. Participants then generated a sequence using bars of different heights to represent relative quantitative or qualitative differences in options. Prior research has suggested that when choosing sequences of unavoidable outcomes, people tend to prefer improving sequences (Loewenstein & Prelec, 1991), and that this may be particularly true when aversive outcomes are involved (Story et al., 2013). Results of Study 2 suggest that this preference for sequences, or saving the best for last, may not be as robust as previously indicated. Furthermore, these results provide support for prior conclusions that different methods may occasion different response patterns (Andrade & Hackenberg, 2012). Using graphical representations of sequences (e.g., bars of different heights), may encourage participants to apply a more superficial level of analysis in which preference for improvement is more likely (Frederick & Loewenstein, 2008). Nevertheless, despite a smaller percentage of participants who saved the best for last in Study 2, the percentage of participants who saved the best for last was highest in the food and noxious categories.

Another noteworthy consideration is that although Andrade and Hackenberg replicated the previous findings of Loewenstein and Prelec (1991, 1993) when using questionnaires, they did not when participants were making repeated choices of sequences of (actual) video clips.

There are several differences across the two preparations that could account for the discrepancy in results and are worthy of consideration and future research. One difference is how the value of the sequence was determined in each preparation: qualitative manipulations of the different hypothetical dinners or outings versus delay manipulations. The category or nature of the reinforcer may also influence the choice, as seen in the current Study 2. As such, is it possible that the use of video clips as reinforcers (whose value or preference may be enhanced by continuity of access; Deleon et al., 2014) may not promote negative time preference. Another difference is the type of outcome (or reinforcer) associated with the choice: hypothetical or real. In Andrade and Hackenberg's (2012) study, when participants were choosing among consumable reinforcers, the results were not consistent with negative time preference. As such, comparisons of participant responding when the sequence outcomes are hypothetical or real is an important area for future research.

Moreover, the results of Study 1a suggest that "saving the best for last" may be a more robust phenomenon than a preference for perfect improvement. Although the percentage of participants who elected to save the best for last was high for all array sizes, that percentage of participants who generated a perfectly improving sequence was significantly lower. Furthermore, there was an interaction effect between the size of the array, and the order of exposure, in that generating perfectly improving sequences for arrays of six, seven, or eight outcomes, was more likely when those where experienced toward the end of the survey. The authors believe this is an area that is worth exploring further, as it may shed light on potential interventions for impulsive behavior. Perhaps, exposing people to sequences of increasing array sizes may promote negative time preference, self-control, or delay of gratification.

Delay discounting research has shown that in choice paradigms where only one outcome is selected from the array, delayed outcomes are generally valued less than those without a delay (Ainslie, 1975). In keeping with this logic, if one were presented with an array of five items and given a chance to access each in a sequence of one's own choosing, one might conclude that the item selected first would be the most preferred. Some readers will recognize this arrangement as that characterized by the multiple-stimulus without replacement (MSWO) preference assessment described by DeLeon and Iwata (1996). It is possible, however, that the most preferred item may not be correctly identified in an MSWO, if individuals choose to save the best for last. Not selecting the highest preferred item may be a concern because lesser preferred items may not engender the same level of performance as more preferred items (Glover et al., 2008; Koehler et al., 2005). There is some scattered evidence to suggest that this phenomenon may occur, as it has been reported in the literature (Soldberg et al., 2007), however, there are no published explicit studies, to date, of the "saving the best for last" phenomenon as it relates to the MSWO arrangement. In fact, this may represent a textbook exemplar of the "file drawer" problem (e.g., Shadish et al., 2016), thus, the prevalence of the phenomenon is unclear. Although the current studies did not evaluate saving the best for last in the MSWO, researchers and clinicians should be aware of the existence of this phenomenon when interpreting the results of an MSWO and determining the assessment's construct validity. Future research should be conducted in this area, to develop specific suggestions for researchers and clinicians.

Although the present studies involved a non-clinical population and hypothetical outcomes, determining the extent of the phenomenon (negative time preference or saving the best for last) was thought to be a foundational precursor to applied explorations of behavior change. The authors view self-reports as *one* method for studying human behavior. Recognizing the value of self-reports and expanding our methods in general may give our work as behavior analysts greater relevance beyond our field (Critchfield & Reed, 2017; Vyse, 2013). Furthermore, the authors believe these initial studies were necessary to determine whether the “saving the best for last” phenomenon is real and whether future research in this area should be pursued. The results of the Study 1 suggest that more than half of all participants (M = 62.88%; range, 56.99% to 67.74%) chose to save the best for last. Therefore, this phenomenon is, in our view, worthy of further study. In Study 1b, participants generated very diverse lists of outcomes, which could explain the reduced percentage of participants who saved the best for last, compared to Study 1a. Consequently, in Study 2, the aim was to study how categorically-different sequences (e.g., those involving schoolwork, food, exercise, leisure, and noxious experiences) influence saving the best for last. We recognize, however, correspondence between verbal behavior and nonverbal behavior does not always align. Thus, to inform our collective interpretation of verbal report preferences for event sequences, the researchers concluded that the next logical step was to compare correspondence of responses to hypothetical or real outcomes. The correspondence between verbal report and nonverbal behavior was the subject of the next portion of this series of studies, described in a concurrent submission.

References

Ainslie, G. (1975). Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin, 82*(4), 463-496. https://doi.org/10.1037/h0076860

American Psychological Association, Policy and Planning Board. (2008). *How technology changes everything (and nothing) in psychology: 2008 annual report of the APA Policy and Planning Board.* https://doi.org/10.1037/a0015888

Andrade, L. F., & Hackenberg, T. D. (2012). Saving the best for last? A cross-species analysis of choices between reinforcer sequences. *Journal of the Experimental Analysis of Behavior, 98*(1), 45-64. https://doi.org/10.1901/jeab.2012.98-45

Ariely, D., & Carmon, Z. (2000). Gestalt characteristics of experiences: The defining features of summarized events. *Journal of Behavioral Decision Making, 13*(2), 191-201. https://doi.org/10.1002/(SICI)1099-0771(200004/06)13:2<191::AID-BDM330>3.0.CO;2-A

Ariely, D., & Zauberman, G. (2000). On the making of an experience: The effects of breaking and combining experiences on their overall evaluation. *Journal of Behavioral Decision Making, 13*(2), 219-232. https://doi.org/10.1002/(SICI)1099-0771(200004/06)13:2<219::AID-BDM331>3.0.CO;2-P

Barney, C. C., Tervo, R., Wilcox, G. L., & Symons, F. J. (2017). A case-controlled investigation of tactile reactivity in young children with and without global developmental delay. *American Journal of Intellectual and Developmental Disabilities, 122*(5), 409-421. https://doi.org/10.1352/1944-7558-122.5.409

Barratt, E. S. (1959). Anxiety and impulsiveness related to psychomotor efficiency. *Perceptual and Motor Skills, 9*(2), 191-198. https://doi.org/10.2466/pms.1959.9.3.191

Chapman, G. B. (1996). Expectations and preferences for sequences of health and money. *Organizational Behavior and Human Decision Processes, 67*(1), 59-75. https://doi.org/10.1006/obhd.1996.0065

Chapman, G. B. (2000). Preferences for improving and declining sequences of health outcomes. *Journal of Behavioral Decision Making, 13*(2), 203-218. https://doi.org/10.1002/(SICI)1099-0771(200004/06)13:2<203::AID-BDM317>3.0.CO;2-S

Cook, J. O., & Barnes, L. W., Jr. (1964). Choice of delay of inevitable shock. *Journal of Abnormal and Social Psychology, 68*(6), 669-672. https://doi.org/10.1037/h0047127

Critchfield, T. S., & Reed, D. D. (2017). The fuzzy concept of applied behavior analysis research. *The Behavior Analyst*, *40*(1), 123-159.

De Wit, H., Flory, J. D., Acheson, A., McCloskey, M. & Manuck, S. B. (2007). IQ and nonplanning impulsivity are independently associated with delay-discounting in middle-aged adults. *Personality and Individual Differences, 42*(1), 111-121. https://doi.org/10.1016/j.paid.2006.06.026

DeLeon, I. G., Chase, J. A., Frank-Crawford, M. A., Carreau-Webster, A. B., Triggs, M. M., Bullock, C. E., & Jennett, H. K. (2014). Distributed and accumulated reinforcement arrangements: Evaluations of efficacy and preference. *Journal of Applied Behavior Analysis, 47*(2), 293–313. https://doi.org/10.1002/jaba.116

DeLeon, I. G., & Iwata, B. A. (1996). Evaluation of a multiple-stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis, 29*(4), 519-533. https://doi.org/10.1901/jaba.1996.29-519

Drolet, A., Lau-Gesk, L., & Scott, C. (2011). The influence of aging on preferences for sequences of mixed affective events. *Journal of Behavioral Decision Making, 24*(3), 293-314. https://doi.org/10.1002/bdm.695

Duckworth, A. L., & Kern, M. L. (2011). A meta-analysis of the convergent validity of self-control measures. *Journal of Research in Personality, 45*(3), 259–268. https://doi.org/10.1016/j.jrp.2011.02.004

Duffy, S., Smith, J., & Woods, K. (2015). How does the preference for increasing payments depend on the size and source of the payments? *Management Science Letters, 5*(12), 1071-1080. https://doi.org/10.5267/j.msl.2015.10.007

Frederick, S., & Loewenstein, G. (2008). Conflicting motives in evaluations of sequences. *Journal of Risk and Uncertainty, 37*(2-3), 221-235. https://doi.org/10.1007/s11166-008-9051-z

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature, 40*(2), 351-401. https://doi.org/10.1257/002205102320161311

Gleick, J. (2000). *Faster: The acceleration of just about everything.* New York, NY: Vintage Books.

Glover, A. C., Roane, H. S., Kadey, H. J., & Grow, L. L. (2008). Preference for reinforcers under progressive- and fixed-ratio schedules: A comparison of single and concurrent arrangements. *Journal of Applied Behavior Analysis, 41*(2), 163–176. https://doi.org/10.1901/jaba.2008.41-163

Guyse, J. L., Keller, L. R., & Eppel, T. (2002). Valuing environmental outcomes: Preferences for constant or improving sequences. *Organizational Behavior and Human Decision Processes, 87*(2), 253–277. https://doi.org/10.1006/obhd.2001.2965

Koehler, L. J., Iwata, B. A., Roscoe, E. M., Rolider, N. U., & O'Steen, L. E. (2005). Effects of stimulus variation on the reinforcing capability of nonpreferred stimuli. *Journal of Applied Behavior Analysis, 38*(4), 469–484. https://doi.org/10.1901/jaba.2005.102-04

Kumar, A., & Gilovich, T. (2016). To do or to have, now or later? The preferred consumption profiles of material and experiential purchases. *Journal of Consumer Psychology, 26*(2), 169-178. https://doi.org/10.1016/j.jcps.2015.06.013

Kumar, A., Killingsworth, M. A., & Gilovich, T. (2014). Waiting for Merlot: Anticipatory consumption of experiential and material purchases. *Psychological Science, 25*(12), 1924-1931. https://doi.org/10.1177/0956797614546556

Legg, A. M., & Sweeny, K. (2014). Do you want the good news or the bad news first? The nature and consequences of news order preferences. *Personality and Social Psychology Bulletin, 40*(12), 279-288. https://doi.org/10.1177/0146167213509113

Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., … Wolff, S. (2009). A brief history of the internet. *Computer Communication Review, 39*(5), 22-31. Retrieved from https://arxiv.org/html/cs/9901011?

Löckenhoff, C. E., Reed, A. E., & Maresca, S. N. (2012). Who saves the best for last? Age differences in preferences for affective sequences. *Psychology and Aging, 27*(4), 840-848. https://doi.org/10.1037/a0028747

Löckenhoff, C. E., Rutt, J. L., Samanez-Larkin, G. R., O'Donoghue, T., & Reyna, V. F. (2019). Preferences for temporal sequences of real outcomes differ across domains but do not vary by age. *Journals of Gerontology: Psychological Sciences, 74*(3), 430-439. https://doi.org/10.1093/geronb/gbx094

Loewenstein, G. (1987). Anticipation and the valuation of delayed consumption. *The Economic Journal, 97*(387), 666-684. https://doi.org/10.2307/2232929

Loewenstein, G., & Prelec, D. (1991). Negative time preference. *The American Economic Review, 81*(2), 347-352.

Loewenstein, G. F., & Prelec, D. (1993). Preferences for sequences of outcomes. *Psychological Review, 100*(1), 91-108. https://doi.org/10.1037/0033-295X.100.1.91

Lucci, C. R. (2013). Time, self, and intertemporal choice. *Frontiers in Neuroscience, 7*(7), 1-5. https://doi.org/10.3389/fnins.2013.00040

Madden, G. J., & Johnson, P. S. (2010). A delay-discounting primer. In G. J. Madden & W. K. Bickel (Eds.), *Impulsivity: The behavioral and neurological science of discounting* (pp. 11-37). Washington, DC: American Psychological Association.

Magen, E., Dweck, C. S., & Gross, J. J. (2008). The hidden-zero effect: Representing a single choice as an extended sequence reduces impulsive choice. *Psychological Science, 19*(7), 648-649. https://doi.org/10.1111/j.1467-9280.2008.02137.x

Marshall, L. L., & Kidd, R. F. (1981). Good news or bad news first? *Social Behavior and Personality, 9*(2), 223-226. https://doi.org/10.2224/sbp.1981.9.2.223

Mitchell, L. A., MacDonald, R. A. R., & Brodie, E. E. (2004). Temperature and the cold pressor test. *The Journal of Pain, 5*(4), 233-238. https://doi.org/10.1016/j.jpain.2004.03.004

Odum, A. L. (2011). Delay discounting: I'm a k, you're a k. *Journal of the Experimental Analysis of Behavior, 96*(3), 427-439. https://doi.org/10.1901/jeab.2011.96-423

Olson, M., & Bailey, M. (1981). Positive time preference. *Journal of Political Economy, 89*(1), 1-25. https://doi.org/10.1086/260947

Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt Impulsiveness Scale. *Journal of Clinical Psychology, 51*(6), 768-774. https://doi.org/10.1002/1097-4679(199511)51:6<768::AID-JCLP2270510607>3.0.CO;2-1

Qualtrics® XM (2018) [Online survey platform]. Provo, UT: Qualtrics.

Shadish, W. R., Zelinsky, N. A. M., Vevea, J. L., & Kratochwill, T. R. (2016). A survey of publication practices of single-case design researchers when treatments have small or large effects. *Journal of Applied Behavior Analysis, 49*(3)*,* 656–673. https://doi.org/10.1002/jaba.308

Sidman, M. (1960). *Tactics of scientific research: Evaluating experimental data in psychology.* New York: Basic Books.

Soldberg, K. M., Hanley, G. P., Layer, S. A., & Ingvarsson, E. Y. (2007). The effects of reinforcer pairing and fading on preschoolers' snack selections. *Journal of Applied Behavior Analysis, 40*(4), 633-644. https://doi.org/10.1901/jaba.2007.633–644

Stanford, M. S., Mathias, C. W., Dougherty, D. M., Lake, S. L. Anderson, N. E., & Patton, J. H. (2009). Fifty years of the Barratt Impulsiveness Scale: An update and review. *Personality and Individual Differences, 47*(5), 385-395. https://doi.org/10.1016/j.paid.2009.04.008

Story, G. W., Vlaev, I., Seymour, B., Winston, J. S., Darzi, A., & Dolan, R. J. (2013). Dread and the disvalue of future pain. *PLoS Computational Biology, 9*(11), 1-18. https://doi.org/10.1371/journal.pcbi.1003335

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481)*,* 453-458. https://doi.org/10.1126/science.7455683 von Baeyer, C. L., Piira, T., Chambers, C. T., Trapanotto, M., & Zelter, L. K. (2005). Guidelines for the cold pressor task as an experimental pain stimulus for use with children. *The Journal of Pain, 6*(4)*,* 218-227. https://doi.org/10.1016/j.jpain.2005.01.349

Vyse, S. (2013). Changing course. *The Behavior Analyst, 36*(1), 123-135. https://doi.org/10.1007/BF03392295

Yarnitsky, D. & Pud, D. (2004). Quantitative sensory testing. In C. D. Binnie, R. Cooper, F. Mauguière, J. W. Osselton, P. F. Prior, & B. M. Tedman (Eds.), *Clinical Neurophysiology (Revised and Enlarged Edition): Vol. 1. EMG, Nerve Conduction and Evoked Potentials* (pp. 309-336). Amsterdam: Elsevier.

Table 1

Study 1 Participant Characteristics

Demographic Characteristics	*f* (%)
Self-identified gender	
Man	55 (19.71)
Woman	223 (79.93)
Gender variant/non-conforming	1 (0.36)
Age	
18-24 years old	237 (84.95)
25-34 years old	27 (9.68)
35-44 years old	10 (3.58)
45-54 years old	5 (1.79)
Major	
Psychology	118 (42.29)
Biological sciences	60 (21.51)
Other	94 (33.69)
No response	7 (2.51)

Note. *f* = frequency of participants.

Table 2

Percentage of Participants who Selected to Save the Best for Last in Each Question of Study 1a

Time frame in choice	Percentage who Saved the Best for Last
1. This coming weekend and the weekend after.	80
2. This coming weekend and in 6 months (26 weeks).	37*
3. In 6 months (26 weeks from now) and the weekend after (27 weeks from now).	68*
4. From 10:00 to 10:25 and 10:30 to 10:55.	82
5. At 9:00 am and at 1:00 pm on the same day.	86
6. On the 21st of the month and on the 22nd of the same month.	81

Note. The table displays the data on the direct replication (lines 1-3) and systematic extension (lines 4-5) of Loewenstein and Prelec (1991, 1993).

* Significantly different from that obtained by Loewenstein and Prelec (1991, 1993), $p < .05$.

Table 3

Percentage of Participants who Generated a Perfectly Improving Sequence for Array Sizes 3 through 8 Separated by Experience-Order Group in Study 1

	Order				
Array Size	Increasing	Decreasing	Random	χ^2 (df=2)	*p*
3	51.0	42.9	39.5	2.393	.302
4	44.3	35.7	27.9	4.410	.110
5	40.2	23.8	37.2	3.970	.137
6*	40.7	21.4	27.9	6.981	.030
7*	39.7	19.0	14.0	14.750	.001
8*	37.6	14.3	20.9	11.363	.003
n	194	42	43		
$\hat{Y}$	-2.30X +54.90	-5.58X +56.87	-4.12X +50.57		
R^2	.797	.924	.639		

Note. χ^2 (df = 2) = Chi-squared tests comparing the percentage of participants who generated a perfectly improving sequence across groups. *p* = test of significance. Percentage of participants who generated a perfectly improving sequence did not differ significantly across groups for array sizes 3 through 5, but did differ significantly for array sizes 6 through 8, as indicated by the asterisk (*).

n = sample size. $\hat{Y}$ = regression equation. R^2 = variance explained by the array size.

Table 4

Frequency (and Percentage) of Participants Who Saved the Best for Last or Generated the Two Most Common Sequence Patterns Per Array Size in Study 1b

Array Size	Saved the Best for Last	Perfectly Improving	Perfectly Worsening
2	189 (67.74)	189 (67.74)	90 (32.26)
3	180 (64.52)	134 (48.03)	55 (19.71)
4	181 (64.87)	113 (40.50)	50 (17.92)
5	181 (64.87)	104 (37.28)	48 (17.20)
6	171 (61.29)	100 (35.84)	51 (18.28)
7	159 (56.99)	91 (32.62)	49 (17.56)
8	167 (59.86)	88 (31.54)	48 (17.20)

Note. Participants who generated a perfectly improving sequence also saved the best for last, but some participants who saved the best for last did not generate perfectly improving sequences for array sizes 3 through 8.

Table 5

Study 2 Participant Characteristics

Demographic Characteristics	*f* (%)
Self-identified gender	
Man	34 (17.71)
Woman	155 (80.73)
Gender variant/non-conforming	1 (0.52)
Not specified	2 (1.04)
Age	
18-24 years old	167 (86.98)
25-34 years old	19 (9.90)
35-44 years old	4 (2.08)
45-54 years old	2 (1.04)
Major	
Psychology	121 (63.02)
Biological sciences	32 (16.67)
Other	39 (20.31)

Note. N = 192. f = frequency of participants.

Table 6

Percentage of Participants who Selected to Save the Best for Last in Each Category of Study 2

Category	Percentage who Saved the Best for Last	χ^2 (1)	*p*
Food	30.73*	13.81	<.001
Noxious stimuli	28.65*	8.97	.003
Exercise	25.52	3.66	.060
Schoolwork	23.44	1.42	.230
Leisure	20.31	0.01	.910
Mixed	16.67	1.33	.250

Note. In only two categories (food and noxious stimuli), * the percentage of participants who saved the best for last was significantly greater than the expected percentage, had participants responded randomly ($p < .05$; exact *p*-values are reported in the *p* column).

Table 7

Study 2. BIS-11 and Subscale Reliability and Mean Scores

	Scale Reliability		Participant Scores	
	N of Items	Cronbach's α	Mean	(SD)
BIS-11	30	.81	60.91	(9.60)
Subscales				
Non-Planning Impulsiveness	11	.65	23.20	(4.37)
Motor Impulsiveness	11	.63	20.68	(3.99)
Attentional Impulsiveness	8	.73	17.03	(3.80)

Note. SD = Standard deviation.

Table 8

Study 2. Mean Spearman's Rho and Pearson Correlations among Impulsiveness and Preference for Sequences for Each Category

			Pearson Correlation (r) Between r_s and			
	Spearman's Rho (r_s)		BIS-11		BIS-11 Non-planning	
Category	*M*	(SD)	*r*	*p*	*r*	*p*
Food	.20	.83	-.07	.36	-.01	.86
Noxious	.24	.79	.00	.95	-.01	.93
Exercise	.30	.78	.02	.81	.16	.03
Schoolwork	.33	.78	-.04	.54	.01	.90
Leisure	.41	.75	-.06	.43	.00	.95
Mixed	-.04	.37	.00	.97	.11	.13

Note. Spearman's Rho (r_s) is the rank-order correlation between preference rank and position in the sequence. *r* = Pearson correlation. *p* = significance.

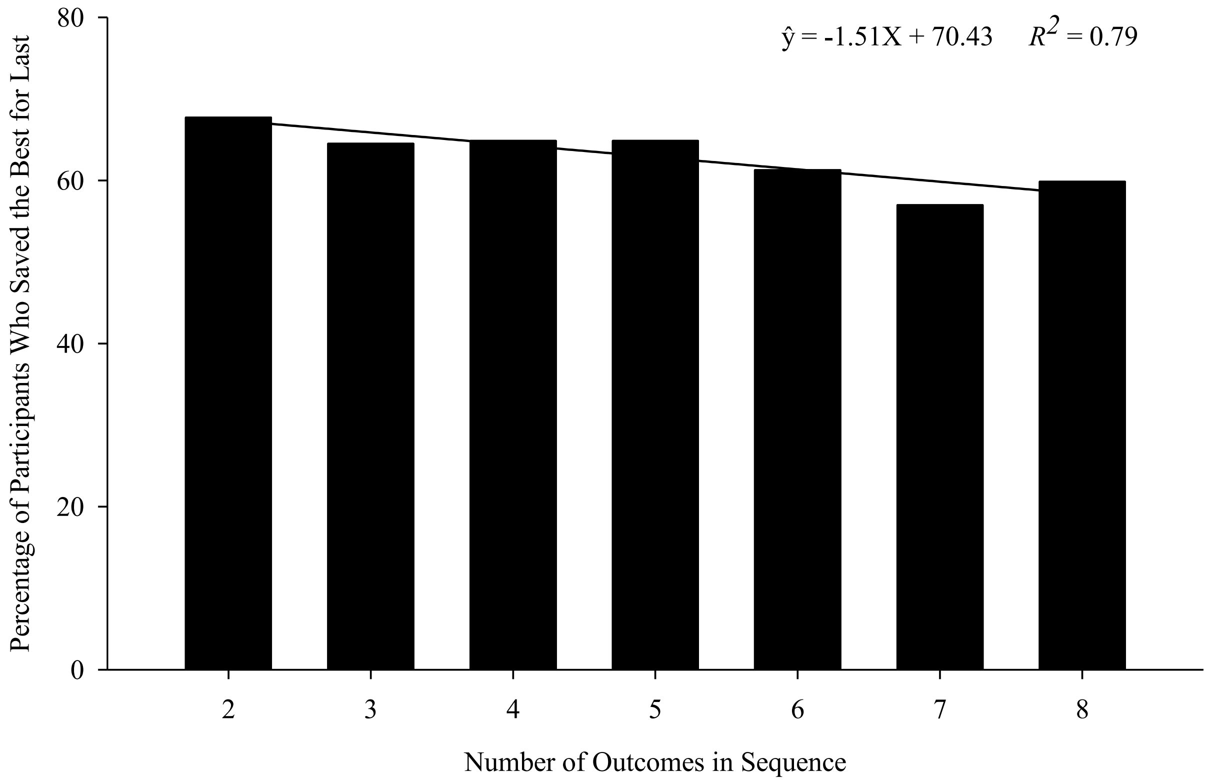

Figure 1. Percentage of participants who saved the best for last for each array size in Study 1b. The percentage of participants who saved the best for last decreased as the array size increased, as indicated by the negative slope of the solid black regression line (B = -1.51, p = .007).

Appendix

Sequence Survey Questions

Leisure

Observing response:

You will now be asked to rank a list of activities in order of preference (1 = most preferred; 5 = least preferred). Before you do so, please complete the following statement:

"I understand I may not like any of these, but if I have to rank them, **<u>of the activities listed</u>**, the activity ranked #1 is my ...

- o MOST favorite (correct answer)
- o LEAST favorite

Q. Rank (drag and drop) the following activities in order of preference (1 = most preferred; 5 = least preferred).

_____ reading a magazine

_____ watching YouTube®

_____ spending time with your phone

_____ playing a paper game like Sudoku or crossword puzzles

_____ playing a computer game

Q. Imagine you get to spend the next couple of hours doing all those activities. In what order would you like to do the following activities?

Select the order in which you would like to experience each activity (first, second, third, fourth, last),

	Reading a Magazine	Watching YouTube®	Spending time with your phone	Playing a paper game like Sudoku or crossword puzzles	Playing a computer game
First	○	○	○	○	○
Second	○	○	○	○	○
Third	○	○	○	○	○
Fourth	○	○	○	○	○
Last	○	○	○	○	○

Schoolwork

Observing response:

You will now be asked to rank a list of activities in order of preference (1 = most preferred; 5 = least preferred). Before you do so, please complete the following statement:

"I understand I may not like any of these, but if I have to rank them, **of the activities listed**, the activity ranked #1 is my ...

- o MOST favorite (correct answer)
- o LEAST favorite

Q. Rank (drag and drop) the following activities in order of preference (1 = most preferred; 5 = least preferred).

_____ Completing worksheets with 3- and 4- digit multiplication problems, without a calculator.

_____ Completing worksheets with long division problems, without a calculator.

_____ Typing handwritten notes onto a computer

_____ Reading a text and answering reading comprehension questions (like those in the SAT)

_____ Work on building a model car with plastic rods and connectors

Q. Imagine you have to spend the next few hours working on school tasks. In what order would you prefer to work on these tasks?

Select the order in which you would like to work on each task (first, second, third, fourth, last),

	Completing worksheets with 3- and 4- digit multiplication problems, without a calculator.	Completing worksheets with long division problems, without a calculator.	Typing handwritten notes onto a computer	Reading a text and answering reading comprehension questions (like those in the SAT)	Work on building a model car with plastic rods and connectors
First	◯	◯	◯	◯	◯
Second	◯	◯	◯	◯	◯
Third	◯	◯	◯	◯	◯
Fourth	◯	◯	◯	◯	◯
Last	◯	◯	◯	◯	◯

Noxious Stimuli

Observing response:

You will now be asked to rank a list of activities in order of preference (1 = most preferred; 5 = least preferred). Before you do so, please complete the following statement:

"I understand I may not like any of these, but if I have to rank them, **of the activities listed**, the activity ranked #1 is my ...

- MOST favorite (correct answer)
- LEAST favorite

Q. Rank (drag and drop) the following activities in order of preference (1 = most preferred; 5 = least preferred).

_____ Submerging your hand in very cold water (1^{o} C, or approximately 33^{o} F)

_____ Having the skin on your inner wrist pricked with a small plastic pin (feels like forcefully touching the sharp end of a pushpin).

_____ Experiencing the equivalent of 4 pounds of pressure in your inner wrist.

_____ Having the end of a thick and stiff nylon thread pressed repeatedly against the skin of your inner wrist 30 times in 30 seconds.

_____ Having a warm probe (50° C, or 122° F) touch the skin of your inner wrist.

Q. Imagine you go to a neurological appointment for testing. You have to experience each of these unpleasant sensations. In what order would you prefer to experience them?

Select the order in which you would like to experience each unpleasant stimulus (first, second, third, fourth, last).

	Submerging your hand in very cold water (1^{o} C, or ~ 33^{o} F).	Having the skin on your inner wrist pricked with a small plastic pin (feels like forcefully touching the sharp end of a pushpin).	Experiencing the equivalent of 4 pounds of pressure in your inner wrist.	Having the end of a thick and stiff nylon thread pressed repeatedly against the skin of your inner wrist 30 times in 30 seconds.	Having a warm probe (50° C, or 122° F) touch the skin of your inner wrist.
First	○	○	○	○	○
Second	○	○	○	○	○
Third	○	○	○	○	○
Fourth	○	○	○	○	○
Last	○	○	○	○	○

Exercise

Observing response:

You will now be asked to rank a list of activities in order of preference (1 = most preferred; 5 = least preferred). Before you do so, please complete the following statement:

"I understand I may not like any of these, but if I have to rank them, **of the activities listed**, the activity ranked #1 is my ...

- o MOST favorite (correct answer)
- o LEAST favorite

Q. Rank (drag and drop) the following activities in order of preference (1 = most preferred; 5 = least preferred).

_____ Doing jumping jacks
_____ Doing sit-ups
_____ Doing push-ups
_____ Running in place
_____ Doing squats

Q. Imagine you have to do all these for 1 minute each, in an exercise session. In what order would you prefer to do these exercises?

Select the order in which you would like to do these exercises (first, second, third, fourth, last).

	Jumping jacks	Sit-ups	Push-ups	Running in place	Squats
First	○	○	○	○	○
Second	○	○	○	○	○
Third	○	○	○	○	○
Fourth	○	○	○	○	○
Last	○	○	○	○	○

Food

Observing response:

You will now be asked to rank a list of foods in order of preference (1 = most preferred; 5 = least preferred). Before you do so, please complete the following statement:

"I understand I may not like any of these, but if I have to rank them, **<u>of the foods listed</u>**, the food ranked #1 is my ...

- o MOST favorite (correct answer)
- o LEAST favorite

Q. Rank (drag and drop) the following foods in order of preference (1 = most preferred; 5 = least preferred).

_____ Strawberries

_____ Grapes

_____ Cucumber

_____ Tomatoes

_____ Chips

Q. Imagine you have to eat all those foods. In what order would you like to eat them?

Select the order in which you would like to eat each food (first, second, third, fourth, last).

	Strawberries	Grapes	Cucumber	Tomatoes	Chips
First	◯	◯	◯	◯	◯
Second	◯	◯	◯	◯	◯
Third	◯	◯	◯	◯	◯
Fourth	◯	◯	◯	◯	◯
Last	◯	◯	◯	◯	◯

Mixed

Observing response:

You will now be asked to rank a list of activities in order of preference (1 = most preferred; 5 = least preferred). Before you do so, please complete the following statement:

"I understand I may not like any of these, but if I have to rank them, **of the activities listed**, the activity ranked #1 is my ...

- o MOST favorite (correct answer)
- o LEAST favorite

Q. Rank (drag and drop) the following activities in order of preference (1 = most preferred; 5 = least preferred).

_____ [Leisure #1]

_____ [Schoolwork #1]

_____ [Noxious stimulus #1]

_____ [Food #1]

_____ [Exercise #1]

Q. Imagine you have to do or eat all the things listed above. In what order would you like to do or eat them?

Select the order in which you would like to do or eat each (first, second, third, fourth, last).

	[Leisure #1]	[Schoolwork #1]	[Noxious stimulus #1]	[Food #1]	[Exercise # 1]
First	○	○	○	○	○
Second	○	○	○	○	○
Third	○	○	○	○	○
Fourth	○	○	○	○	○
Last	○	○	○	○	○