Framing attributes in similarity judgments and intertemporal choices

Jeffrey R. Stevens[1], Francine W. Goh[1], & Tyler Cully[1]

[1] University of Nebraska-Lincoln

Similarity models provide an alternative approach to intertemporal choice. Instead of calculating an overall value for options, decision makers compare the similarity of option attributes and make a decision based on similarity. Similarity judgments for reward amounts and time delays depend on both the numerical difference ($x_2 - x_1$) and ratio (x_1/x_2) of quantitative values. Changing units of these attribute values (e.g., days vs. weeks) can alter the numerical difference while maintaining the ratio. For example, framing a pair of delays in the unit of weeks (1 vs. 2) or days (7 vs. 14) both result in a ratio of 1/2. Yet the numerical difference between the delays differs depending on the unit (1 for weeks and 7 for days). Here we had participants make similarity judgments and intertemporal choices with amounts framed as dollars or cents and delays framed as days or weeks. We predicted that they units of amounts and delays would influence similarity judgments which would then influence intertemporal choices. We found that participants judged amounts framed as cents as less similar than dollars, and this resulted in more patient intertemporal choices. Additionally, they judged delays framed as weeks as more similar than days, but the framing did not influence choice. These findings suggest that the units in which amounts and delays are framed can influence their similarity judgments, which can shape intertemporal choices. These unit effects may guide stakeholders in framing aspects of intertemporal choices in different units to nudge decision makers into either more impulsive or patient choice.

Keywords: framing, intertemporal choice, similarity judgments, unit effect
Word count: 7500

Introduction

From choosing between healthy and unhealthy foods to deciding whether to invest in retirement, we frequently face important intertemporal choices that involve benefits available at different times in the future. Each option comprises a reward amount and a time delay to receive the reward. The dominant theoretical framework for understanding these decisions (temporal discounting) involves calculating the discounted value of options, that is, the value of the reward discounted by the time delay to receiving the reward (Doyle, 2013; Regenwetter et al., 2018). Decision makers then compare the discounted values to make a choice.

Similarity models instead focus on comparing decision attributes (reward amounts, time delays) rather than options (Goh & Stevens, 2021). In similarity models of intertemporal choice, decision makers compare the similarity of the benefit attribute reward amounts and the similarity of time delays using a decision tree. If one attribute is similar and the other dissimilar, decision makers will use the dissimilar attribute to make the choice (Rubinstein, 1988; Leland, 2002). For instance, for the choice between \$5 in 2 days and \$7 in 14 days, \$5 and \$7 may be considered similar, whereas 2 days and 14 days may be dissimilar. Therefore, a decision maker may ignore the similar reward amounts, focus on the dissimilar time delays, and choose the sooner option. Though the similarity models faces limits to its application, decision tree models have shown superior predictive ability in a variety of domains (e.g., Luan et al., 2011; Delgado-Gomez et al., 2016; Morris & Perna, 2018), and similarity models can predict choice quite well, outperforming discounting models of intertemporal choice (Stevens, 2016).

Models of intertemporal choice are important because they can predict cognitive processes in these decisions. Because intertemporal choices are implicated in critical parts of individual and societal well-being—including physical and mental health (Story et al., 2014; Amlung et al., 2017; Bickel et al., 2019), financial decisions (Kim & McKinnon, 2020), and environmental sustainability (Hardisty & Weber, 2009)—understanding the cognitive processes of these choices can allow us to frame them in a way that may nudge people into making better decisions. For example, fuzzy trace theory and query theory predict that making certain attributes more

Version: 2021-09-03
Jeffrey R. Stevens, https://orcid.org/0000-0003-2375-1360.
Francine W. Goh, https://orcid.org/0000-0002-7364-4398.
Correspondence concerning this article should be addressed to Jeffrey R. Stevens, B83 East Stadium, Department of Psychology, Center for Brain, Biology & Behavior, University of Nebraska-Lincoln, Lincoln, Nebraska 68588, USA. E-mail: jeffrey.r.stevens@gmail.com

salient or presenting attributes in a particular order can influence cognitive processing and thus choice (Weber et al., 2007; Rahimi-Golkhandan et al., 2017). Here, we argue that varying the numerical framing of attributes of intertemporal choices can alter similarity judgments, thereby influencing intertemporal choice.

Numerical framing

The way numerical values are framed affects choice. For instance, consumers' perceptions of products can vary depending on whether price reductions are framed in currency amounts or percentage terms (e.g., save \$10 vs. save 10% of \$100, Chen et al., 1998), when product attributes are presented in small compared to large ratio frames (e.g., there is a smaller perceived difference between 99% and 99.7% vs. 0.3% and 1%, Kwong & Wong, 2006), and by the separation of the thousands digit from the hundreds digit in four-digit dollar prices (e.g., \$1493 vs. \$1,493, Coulter et al., 2012). In risky choice, presenting risk statistics on a larger scale (e.g., 1,286 cases out of 10,000 people) compared to a smaller scale (e.g., 24.14 cases out of 100 people) can lead to increased risk perception even when the actual risk is lower (Yamagishi, 1997). Similarly, people prefer lotteries that present the chance of winning on a larger scale (e.g., 9 out of 100 chances) compared to when it is presented on a smaller scale (e.g., 1 out of 10 chances) even if the chances are lower (Pacini & Epstein, 1999). Such numerical framing effects may occur due to the so-called *unit effect*, where differences between values are perceived as larger for quantitative information presented on a scale with many units compared to a scale with fewer units (Burson et al., 2009; Pandelaere et al., 2011; Camilleri & Larrick, 2014; Cadario et al., 2016; Skylark et al., 2021). People tend to focus on numeric information while ignoring unit information (Shen & Urminsky, 2013; Schley et al., 2017). The unit effect is related to the numerosity heuristic, in which people infer higher counts of an item to mean greater magnitude (Pelham et al., 1994; Bagchi & Davis, 2016).

Though the unit effect has been applied primarily to monetary amounts, analogous effects occur with time delays. Temporal framing research has shown that people prefer a cost framed in days compared to months and a cost framed in months compared to years (Gourville, 2003). This unit framing effect generalizes across a variety of contexts such as car leases, meal delivery services, and savings programs (Goldstein et al., 2016; Atlas & Bartels, 2018; Hershfield et al., 2020). In intertemporal choice, changing the way dates are framed affects choice. The date/delay effect suggests that people choose the larger, later option more when time is presented in calendar date format compared to delay period (Read et al., 2005; LeBoeuf, 2006; DeHart & Odum, 2015). Additionally, people prefer larger, later options when delays to hedonic rewards are expressed in larger compared to smaller units (e.g., days versus hours, Siddiqui et al., 2018).

We propose that the unit effect observed in both monetary amounts and time delays may be attributable to similarity judgments. In Rubinstein's (1988) original formulation of the similarity model, he suggested two key relationships within attribute values that may influence similarity judgments: numerical differences ($x_2 - x_1$) and numerical ratios (x_1/x_2). Though Buschena and Zilberman (1999) demonstrated that only differences influence similarity judgments of probabilities in risky choice, Stevens and Soh (2018) showed that both difference and ratio independently drive similarity judgments for reward amounts and time delays in intertemporal choice. Thus, if ratio is held constant, changes in numerical difference influence similarity judgments. For example, though 1 vs. 3 and 100 vs. 300 have the same ratio of 1/3, the difference of 200 results in 100 vs. 300 being judged as less similar than 1 vs. 3. Changing units from cents to dollars or days to weeks results in maintaining consistent numerical ratios but changing differences. We suggest that the changes in numerical difference alters similarity judgments, which have downstream effects on choice.

Present study

The primary aim of the present study was to investigate how the framing of reward amounts and time delays can influence people's similarity judgments, which can in turn shape their intertemporal choices. To address this question, we framed reward amounts and time delays in different units and measured participants' similarity judgments and intertemporal choices. Specifically, we framed rewards as dollars or cents and delays as days or weeks. Shifting rewards from dollars to cents will increase the numerical difference, which should decrease amount similarity judgments for cents, thereby inducing preferences for the larger, later option. Shifting delays from days to weeks will decrease the numerical difference, which should increase delay similarity judgments for weeks, again inducing stronger preferences for the larger, later option. Therefore, cognitive processes generating similarity judgments may offer a mechanism that results in the unit effect.

Another aim of this study was to investigate whether an individual's mathematical abilities would moderate the relationship between unit frames and judgments and choices. We assessed participant statistical numeracy (understanding of statistical and probabilistic computations, Cokely et al., 2012). More numerate individuals are less influenced by choice framing effects, can better evaluate everyday risks, and tend to prefer delayed but larger rewards in intertemporal choice (Peters, 2012; Ghazal et al., 2014). We also assessed number line estimation accuracy (representations of numerical magnitude, Peters et al., 2008) because it has a unique relationship with decision making, above and beyond statistical numeracy (Park & Cho, 2019). Therefore, we predicted that individuals with higher levels of numeracy and more accurate estimation should be less susceptible to unit framing.

To investigate these aims, we conducted a series of experiments on similarity judgments and intertemporal choices. The first set of experiments varied whether reward amounts were framed as dollars or cents to explore effects on amount similarity judgments and intertemporal choices. The second set of experiments varied whether time delays were framed as days or weeks to explore effects of delay similarity judgments and intertemporal choices.

Amount similarity in intertemporal choice

In the first set of experiments, we framed reward amounts as dollars and cents (100 cents in every dollar). Because cents involve larger numerals due to more individual units, we predicted that, compared to participants in the dollars condition, participants in the cents condition would (1) judge reward amounts as less similar. These similarity judgments should in turn drive participants to (2) choose the larger, later options more often, thereby producing the unit effect on choice. Further, we predicted that (3) amount similarity judgments would mediate the relationship between the frames and the intertemporal choice supporting similarity judgments as a mechanism of the unit effect. Finally, we predicted that (4) numeracy and numerical estimation errors would moderate the unit effect, with highly numerate participants showing weaker effects.

Methods

We collected data on amount similarity judgments in two data sets with independent sets of participants.

Participants. For the first data set, we recruited a sample of participants from the University of Nebraska-Lincoln (UNL) Department of Psychology subject pool between Nov 2017 - Mar 2020. We collected data from 234 undergraduate students in total. We excluded 48 participants who failed attention checks (e.g., rated identical values as dissimilar, rated very different values as similar, or chose the later intertemporal choice option when the amounts were identical), leaving 186 participants. We used sequential hypothesis testing with Bayes factors (Schönbrodt et al., 2017) to determine the final sample size (see Data analysis section). This resulted in a final sample size of 112 participant who were on average 19.8 (range = 18-26) years of age and of whom 74 (66.1%) identified as female, 38 (33.9%) identified as male, and 0 (0%) identified as neither male nor female (full demographics in Table S1).

For the second data set, we recruited a new sample of participants from the same subject pool between Mar 2020 - Oct 2020. We collected data from 259 undergraduate students in total. We excluded 59 participants who failed attention checks, leaving 200 participants. After sequential hypothesis testing, the final sample size was 98 participants who were on average 20.0 (range = 18-38) years of age and of whom 70 (71.4%) identified as female, 26 (26.5%) identified as male, and 2 (2%) identified as neither male nor female (full demographics in Table S1).

Participants in both studies received research credit for their participation. All procedures were conducted in an ethical and responsible manner, in full compliance with all relevant codes of experimentation and legislation and were approved by the UNL Internal Review Board (protocol # 13118). All participants gave consent to participate, and they acknowledged that de-identified data could be published publicly.

Procedures. For data set 1, participants experienced the experiment on a desktop computer in a private room with at most one other participant in the room and the experimenter outside of the room. The experimental stimuli were presented using PsychoPy (Peirce et al., 2019). For data set 2, the experiment switched to an online study on Qualtrics that the participants could complete on their own in whatever location they chose.

Participants provided responses to intertemporal choices, binary choice amount similarity judgments, slider-based amount similarity judgments, binary choice delay similarity judgments, slider-based delay similarity judgments, numeracy questions, slider-based number line estimation questions, and demographics. Intertemporal choices involved 34 hypothetical choices between smaller, sooner and larger, later payoffs (e.g., "Which would you prefer to receive? 11 dollars in 7 days or 13 dollars in 14 days?") (Table S2). Participants in both data sets experienced the same intertemporal choices in a random order.

Binary amount similarity judgments involved 34 judgments between two monetary amounts (e.g., Do you consider these values to be similar or dissimilar? 11 dollars and 13 dollars) using all amount pairs in Table S2 in a random order. Binary delay similarity judgments were comparable to the amount judgments but adapted for delays (e.g., 7 days vs. 14 days; Table S3). In addition to judgments between amount and delay pairs, participants made slider-based judgments. For these judgments, participants were instructed to "Choose the smallest value that you would consider to be dissimilar from [reference value]." They then were presented with a slider that ranged from the reference value + 1 to the reference value + 7. For example, if the reference amount were 3 dollars, they would see 4, 5, 6, 7, 8, 9, and 10 on the slider. Both amounts and delays used the following reference values: 3, 5, 6, 8, 10, 12, 14, 16. Slider-based data were not analyzed for this study.

We assessed numeracy with the four multiple choice questions from the Berlin Numeracy Test (Cokely et al., 2012) and scored them as correct or incorrect. As an additional measure of numerical skill, we assessed errors with 22 questions from a mental number line task (Peters et al., 2008), which included a slider ranging from 1 to 1000 (initially anchored at 500). We instructed participants to "Click on the number line between 1 and 1000 where you think the number [reference value] falls." Participants experienced the following reference values in random order: 2, 5, 18, 34, 56, 78, 100, 122, 147, 150, 163, 179, 246, 366, 486, 606, 722, 725, 738, 754, 818, 938. We subtracted the response from the reference value and took the absolute value to measure accuracy. Participants then provided their age, gender, racial/ethnic identity, university major, and parental income level.

Experimental conditions. Monetary amounts were framed in either US dollars or cents. So for all intertemporal choices and amount similarity judgments, participants were randomly assigned to view the monetary amounts as dollars (e.g., "11 dollars") or the equivalent number of cents (e.g., "1100 cents"). The words "dollars" and "cents" were also used in the instructions any time monetary amounts were referenced. Time delays were always in days.

Data analysis. We analyzed data from the project using R [Version 4.1.1; R Core Team (2017)] and the R-packages *BayesFactor* [Version 0.9.12.4.2; Morey and Rouder (2015)], *bayestestR* [Version 0.10.5; Makowski et al. (2019)], *ggdist* [Version 3.0.0; Kay (2021)], *here* [Version 1.0.1; Müller (2017)], *lme4* [Version 1.1.27.1; Bates et al. (2015)], *mediation* [Version 4.5.0; Tingley et al. (2014); Imai et al. (2010b); Imai et al. (2010a); Imai et al. (2011); Imai and Yamamoto (2013)], *papaja* [Version 0.1.0.9997; Aust and Barth (2017)], *performance* [Version 0.7.3; Lüdecke et al. (2020)], and *tidyverse* [Version 1.3.1; Wickham et al. (2019)]. Data, analysis scripts, supplementary materials, and the reproducible research materials are available at the Open Science Framework (https://osf.io/xnwra/).

We calculated Bayes factors (BF_{10}) to provide the weight of evidence for the alternative hypothesis relative to the null hypothesis (Wagenmakers, 2007). For example, BF_{10} = 10 means that the evidence for the alternative hypothesis is 10 times stronger than the evidence for the null hypothesis. Bayes factors between 1-3 provide only anecdotal evidence, those between 3-10 provides moderate evidence, those between 10-100 provide strong evidence, and those above 100 provide very strong evidence for the alternative over the null hypothesis (Andraszewicz et al., 2015). Bayes factors associated with generalized linear mixed models were converted from Bayesian Information Criterion (BIC) using $BF_{10} = e^{\frac{BIC_{null}-BIC_{alternative}}{2}}$ (Wagenmakers, 2007).

The sample size for both studies was determined by an optional stopping rule based on Bayes factors (Schönbrodt et al., 2017). We calculated the Bayes factors for the t-tests between conditions for the binary amount similarity judgments and the intertemporal choices and stopped analysis when both reached either 0.33 or 3. We analyzed the first 40 participants in each condition and sequentially added pairs of participants (one for each condition) in order of testing until our Bayes factor thresholds were met.

To compare similarity judgments and intertemporal choices between the two framing conditions, we computed frequentist and Bayes factor t-tests. For the Bayes factor t-tests, we used the `ttestBF()` function from the *BayesFactor* package (Morey & Rouder, 2015) using the default settings for the priors (Cauchy distributions for effect sizes and noninformative/uniform distributions for variance, Rouder et al., 2009). Between-subjects 95% confidence intervals are presented in brackets after parameter estimates.

For mediation analysis, we used the `mediate()` function from the *mediation* package (Tingley et al., 2014). We first computed two generalized linear mixed models using the `glmer()` function from the *lme4* package (Bates et al., 2015). The first model had the similarity judgment (similar = 1, dissimilar = 0) as the response variable, condition (dollars or cents) as a fixed effect, participant as a random effect, and a binomial error distribution. The second model had the intertemporal choice as the response variable (larger, later = 1, smaller, sooner = 0), condition and the similarity judgment as fixed effects (no interaction included), participant as a random effect, and a binomial error distribution. The models were conducted at the individual trial level, so condition, similarity judgment, and intertemporal choice were all binary variables. We inputted these models into the `mediate()` function to calculate the average causal mediation effect, average direct effect, and total effect. We used the default quasi-Bayesian approximation for 95% confidence intervals for effects.

To assess moderation effects of numerical ability on similarity judgments and intertemporal choice, we conducted model selection procedures comparing the intercept only model, numerical ability model (numeracy or number line estimation errors), condition model (dollars/cents or delays/weeks), condition and numerical ability model without an interaction, and condition and numerical ability model with an interaction. We calculated Bayes factors for each model with the intercept only model as the null model using the BIC estimation procedure.

Results

Our first hypothesis predicted that, compared to participants in the dollars condition, participants in the cents condition would judge reward amounts as less similar. In data set 1, participants in the cents condition judged 44.2% [37.4, 51.0] of the questions as similar, whereas participants in the dollars condition judged 57.4% [49.2, 65.5] as similar (Figure 1a). This provided moderate evidence of fewer similarity judgments in the cents condition (ΔM = 13.21, 95% CI [2.73, 23.70], t(106.69) = 2.50, p = .014, BF_{10} = 3.16). In

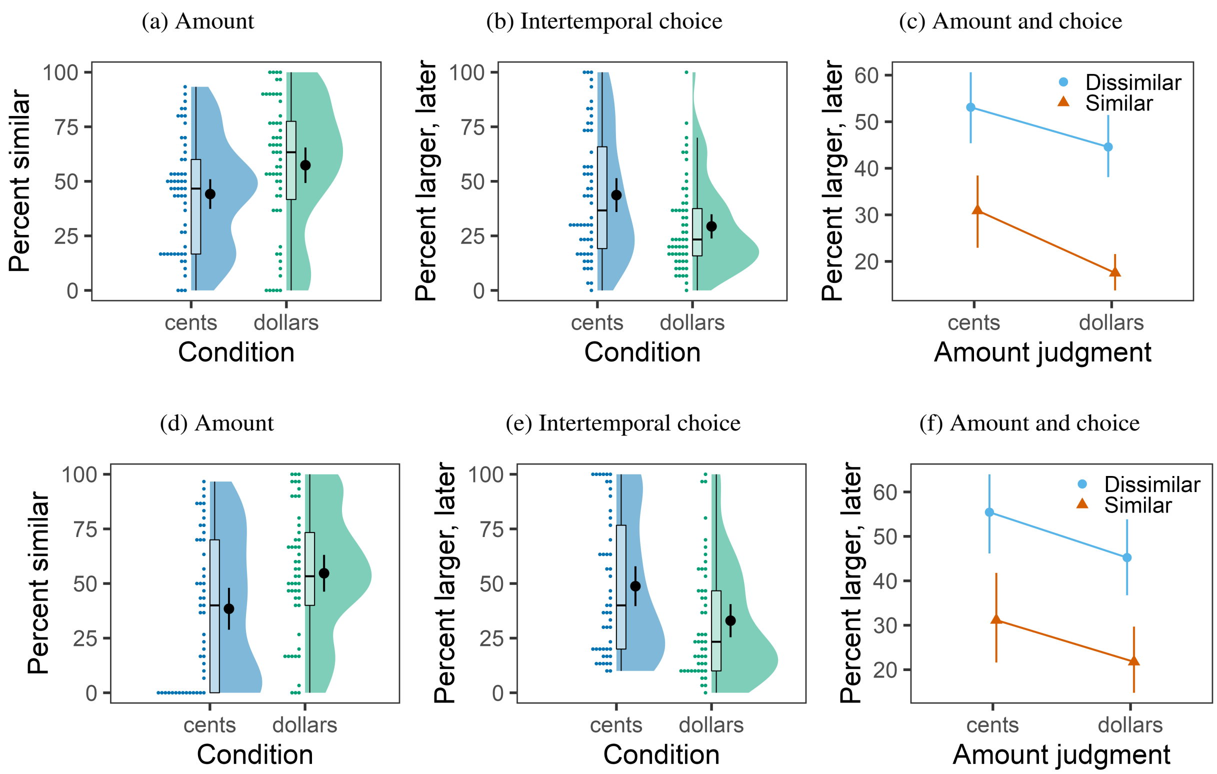

Figure 1. Amount framing similarity judgments, intertemporal choices, and interactions for cents vs. dollars conditions for (a-c) data set 1 (N=56 per condition) and (d-f) data set 2 (N=49 per condition). (a, d) Participants judged pairs of reward amounts as similar or dissimilar, and we calculated the mean percent similarity rating over all pairs for each participant. (b, e) Participants choose between smaller, sooner and larger, later options, and we calculated the mean percent choice for the larger, later per participant. (c, f) We calculated the percent choice for the larger, latter option conditional on whether participants judged the amount pair in the choice as similar or dissimilar for each condition. Small dots represent individual means for participants, filled circles/triangles and error bars represent condition means and 95% between-subjects confidence intervals, boxes represent condition interquartile ranges, horizontal lines represent condition medians, and whiskers represent 1.5 times the interquartile range.

data set 2, participants in the cents condition judged 38.4% [28.9, 48.0] of the questions as similar, whereas participants in the dollars condition judged 54.7% [46.3, 63.1] as similar (Figure 1d). This provided moderate evidence of fewer similarity judgments in the cents condition (ΔM = 16.26, 95% CI [3.67, 28.85], $t(94.43) = 2.56$, $p = .012$, $\text{BF}_{10} = 3.72$). Thus, framing condition influenced amount similarity judgments in the predicted direction in both data sets.

Our second hypothesis predicted that, compared to participants in the dollars condition, participants in the cents condition would choose the larger, later option more often. In data set 1, participants in the cents condition chose larger, later options 43.7% [35.9, 51.4] of the time, whereas participants in the dollars condition chose 29.3% [23.8, 34.9] larger, later (Figure 1b). This provided strong evidence of fewer similarity judgments in the cents condition (ΔM = 14.35, 95% CI [4.91, 23.78], $t(99.70) = 3.02$, $p = .003$, $\text{BF}_{10} = 10.79$). In data set 2, participants in the cents condition chose larger, later options 48.8% [39.7, 57.9] of the time, whereas participants in the dollars condition chose 33.0% [25.4, 40.6] larger, later (Figure 1e). This provided moderate evidence of fewer similarity judgments in the cents condition (ΔM = 15.78, 95% CI [4.07, 27.49], $t(92.91) = 2.68$, $p = .009$, $\text{BF}_{10} = 4.79$). Thus, framing condition influenced intertemporal choices in the predicted direction in both data sets.

Our third hypothesis predicted that amount similarity judgments would mediate the relationship between the framing condition and intertemporal choice (Figures 1c&f). In data set 1, the causal mediation analysis found a direct effect of -0.11 [-0.19, -0.02] ($p = .014$) and a mediation effect of -0.03 [-0.06, -0.01] ($p = .014$), which accounted for 23.0% [4.5, 63.2] of the total effect (-0.14 [-0.23, -0.05], $p = .006$). In data set 2, the causal mediation analysis found a direct effect of -0.13 [-0.25, -0.02] ($p = .034$) and a mediation effect of -0.05 [-0.08, -0.02] ($p = .008$), which accounted for 25.9% [9.1, 76.1] of the total effect (-0.18 [-0.29, -0.06], $p = < .001$). Thus, amount similarity judgments partially mediated the effect of framing on intertemporal choice in both data sets.

Our fourth hypothesis predicted that numeracy and number line estimation errors would moderate the relationship between the framing condition and both amount similarity judgments and intertemporal choice. In data set 1, there was strong evidence that numeracy and number line estimate errors (log-transformed) negatively correlated ($r = -.31$, 95% CI $[-.47, -.12]$, $t(102) = -3.24$, $p = .002$, $BF_{10} = 25.34$). Model selection procedures did not favor the interaction model with condition and numeracy or number line error for either similarity judgments or intertemporal choice (Table S5). In data set 2, there was moderate evidence that numeracy and number line estimate errors were not correlated ($r = .02$, 95% CI $[-.18, .22]$, $t(92) = 0.19$, $p = .853$, $BF_{10} = 0.24$). Model selection procedures did not favor the interaction model with condition and numeracy or number line errors for either similarity judgments or intertemporal choice (Table S5). Thus, numerical ability did not moderate the effect of condition on amount similarity judgments or intertemporal choices.

Summary

In two data sets, participants made amount similarity judgments and intertemporal choices with the reward amounts framed as either dollars or cents. On average, participants in the dollars condition rated reward amount pairs as more similar than participants in the cents condition. This finding supports the notion that, though both frames have the same numerical ratio between reward amounts, the cents condition results in a larger numerical difference, which resulted in those reward amounts being perceived as more dissimilar than dollars, even though the ratios and monetary values are identical. Participants in the cents condition also chose the larger, later option more frequently than those in the dollars condition. This follows from their amount similarity judgments. Participants who judge amounts as less similar should focus more on amounts compared to delays when making intertemporal choices, which would result in choosing the larger, later option more frequently. Thus, the difference between conditions aligns with the similarity account of intertemporal choice. To more formally test whether the similarity judgments mediated the relationship between framing condition and intertemporal choice, we conducted a mediation analysis. This analysis, in fact, demonstrated that similarity judgments partially mediated the effect of framing condition on choice. Finally, we investigated whether numerical abilities (as assessed by numeracy and number line estimation errors) moderated the effect of framing condition on amount similarity judgments and intertemporal choice. Though the two measures of numerical ability correlated in one data set, neither moderated the effect of framing condition on similarity judgments or choice.

Delay similarity in intertemporal choice

Our second experiment mirrored the research questions and methods of the first experiment, focusing on the time delay component of intertemporal choice. For this experiment, we framed time delays as days and weeks. Because days involve larger numerals, we predicted that, compared to participants in the weeks condition, participants in the days condition would (1) judge time delays as less similar and therefore (2) choose the larger, later options more often. Further, we predicted that (3) delay similarity judgments would mediate the relationship between the frames and intertemporal choice. Finally, we predicted that (4) numeracy and numerical estimation errors would moderate the unit framing effect, with highly numerate participants showing weaker effects.

Methods

We collected data on delay similarity judgments in three data sets with independent sets of participants. The first two data sets paralleled those from the amount experiment with experimentally naive participants from the UNL subject pool. Due to slightly different findings between these two data sets, we replicated them in an Amazon Mechanical Turk population for the third data set.

Participants. For the first data set, we recruited a sample of participants from the UNL Department of Psychology subject pool between Feb 2018 - Mar 2020. We collected data from 247 undergraduate students in total. We excluded 41 participants who failed attention checks, leaving 206 participants. We used sequential hypothesis testing with Bayes factors to determine the final sample size. This resulted in a final sample size of 172 participants who were on average 19.8 (range = 17-26) years of age and of whom 128 (74.4%) identified as female, 43 (25%) identified as male, and 1 (0.6%) identified as neither male nor female (full demographics in Table S4).

For the second data set, we recruited a new sample of participants from the same subject pool between Mar 2020 - Apr 2020. We collected data from 236 undergraduate students in total. We excluded 50 participants who failed attention checks, leaving 186 participants. After sequential hypothesis testing, the final sample size was 92 participants who were on average 19.8 (range = 18-25) years of age and of whom 67 (72.8%) identified as female, 25 (27.2%) identified as male, and 0 (0%) identified as neither male nor female (full demographics in Table S4).

For the third data set, we recruited a sample of participants from Amazon's Mechanical Turk (MTurk) from the United States in Feb 2021. We collected data from 229 participants in total. We excluded 73 participants who failed attention checks, leaving 156 participants. After sequential hypothesis testing, the final sample size was 128 participants who were on average 38.1 (range = 21-70) years of age and of whom 68 (53.1%) identified as female, 60 (46.9%) identified as male, and 0 (0%) identified as neither male nor female (full demographics in Table S4).

Participants in the UNL studies received research credit for their participation, whereas those in MTurk received $1.50. The study took about 15-20 minutes for the UNL students and 10-15 minutes for the MTurk participants.

Procedures. For data set 1, participants experienced the experiment on a desktop computer in a private room with at most one other participant in the room and the experimenter outside of the room. The experimental stimuli were presented using PsychoPy. For data sets 2 and 3, the experiment switched to an online study on Qualtrics that the participants could complete on their own in whatever location they chose.

Participants provided responses to the same tasks and questions as the amount similarity experiments: intertemporal choices, binary choice delay similarity judgments, slider-based delay similarity judgments, binary choice amount similarity judgments, slider-based amount similarity judgments, numeracy questions, slider-based number line estimation questions, and demographics. The only difference in procedure is that the delay similarity tasks preceded the amount similarity tasks in these experiments. Also, MTurk participants only experienced intertemporal choices, binary choice delay similarity judgments, numeracy questions, and demographics (no slider-based similarity judgments or number line estimations).

Experimental conditions. Time delays were framed in either days or weeks. So for all intertemporal choices and delay similarity judgments, participants were randomly assigned to view the time delays as days (e.g., "7 days") or the equivalent number of weeks (e.g., "1 week"). The words "days" and "weeks" were also used in the instructions any time delays were referenced. Amounts were always in dollars.

Data Analysis. We used the same data analysis plan as we used for the amount experiment.

Results

Our first hypothesis predicted that, compared to participants in the weeks condition, participants in the days condition would judge time delays as less similar. In data set 1, participants in the days condition judged 40.0% [35.1, 44.8] of the questions as similar, whereas participants in the weeks condition judged 60.1% [55.3, 64.8] as similar (Figure 2a). This provided extreme evidence of fewer similarity judgments in the days condition (ΔM = 20.07, 95% CI $[13.35, 26.80]$, $t(169.92) = 5.89$, $p < .001$, $BF_{10} = 5.34 \times 10^5$). In data set 2, participants in the days condition judged 35.2% [29.1, 41.2] of the questions as similar, whereas participants in the weeks condition judged 50.3% [44.0, 56.6] as similar (Figure 2d). This provided very strong evidence of fewer similarity judgments in the days condition (ΔM = 15.12, 95% CI $[6.47, 23.78]$, $t(89.85) = 3.47$, $p = .001$, $BF_{10} = 36.37$). In data set 3, participants in the days condition judged 54.7% [48.6, 60.8] of the questions as similar, whereas participants in the weeks condition judged 59.7% [52.1, 67.3] as similar (Figure 2g). This provided moderate evidence of no effect of condition on similarity judgments (ΔM = 5.03, 95% CI $[-4.63, 14.68]$, $t(120.43) = 1.03$, $p = .305$, $BF_{10} = 0.31$). Thus, framing condition influenced delay similarity judgments in the predicted direction in two of the three data sets.

Our second hypothesis predicted that, compared to participants in the weeks condition, participants in the days condition would choose the larger, later option more often. In data set 1, participants in the days condition chose larger, later options 28.8% [23.9, 33.7] of the time, whereas participants in the weeks condition chose 37.9% [33.1, 42.8] larger, later (Figure 2b). This provided moderate evidence of fewer similarity judgments in the days condition (ΔM = 9.15, 95% CI $[2.30, 15.99]$, $t(169.98) = 2.64$, $p = .009$, $BF_{10} = 4.00$). In data set 2, participants in the days condition chose larger, later options 34.4% [26.2, 42.7] of the time, whereas participants in the weeks condition chose 38.3% [30.4, 46.3] larger, later (Figure 2e). This provided moderate evidence of no effect of condition on similarity judgments (ΔM = 3.91, 95% CI $[-7.37, 15.20]$, $t(89.89) = 0.69$, $p = .493$, $BF_{10} = 0.27$). In data set 3, participants in the days condition chose larger, later options 47.8% [38.3, 57.3] of the time, whereas participants in the weeks condition chose 55.2% [45.7, 64.7] larger, later (Figure 2h). This provided moderate evidence of no effect of condition on similarity judgments (ΔM = 7.45, 95% CI $[-5.87, 20.77]$, $t(126.00) = 1.11$, $p = .271$, $BF_{10} = 0.33$). Thus, framing condition influenced intertemporal choices in the predicted direction in one of three data sets.

Our third hypothesis predicted that delay similarity judgments would mediate the relationship between the framing condition and intertemporal choice (Figures 2c,f,i). In data set 1, the causal mediation analysis found a direct effect of 0.13 [0.06, 0.20] (p = < .001) and a mediation effect of -0.03 [-0.04, -0.02] (p = < .001), which accounted for 25.6% [79.6, 11.3] of the total effect (0.11 [0.03, 0.18], p = .004). In data set 2, the causal mediation analysis found a mediation effect of -0.02 [-0.04, 0.00] (p = .002) but no direct effect (0.06 [-0.05, 0.18] (p = .332) or total effect (0.04 [-0.08, 0.16], p = .530). In data set 3, the causal mediation analysis found a mediation effect of -0.02 [-0.04, -0.01] (p = < .001) but no direct effect (0.08 [-0.05, 0.22] (p = .274) or total effect (0.06 [-0.08, 0.21], p = .412). Thus, delay similarity judgments partially mediated the effect of framing condition on intertemporal choice in one data set, but the other two data sets showed no total effects.

Our fourth hypothesis predicted that numeracy and number line estimation errors would moderate the relationship between the framing condition and both delay similarity judgments and intertemporal choice. In data set 1, there was anecdotal evidence that numeracy and number line estimate errors negatively correlated ($r = -.19$, 95% CI $[-.34, -.03]$, $t(152) = -2.40$, $p = .018$, $BF_{10} = 2.84$). Model selection procedures did not favor the interaction model with condition and numeracy or number line errors for either similarity judgments or intertemporal choice (Table S5). In data set 2, there

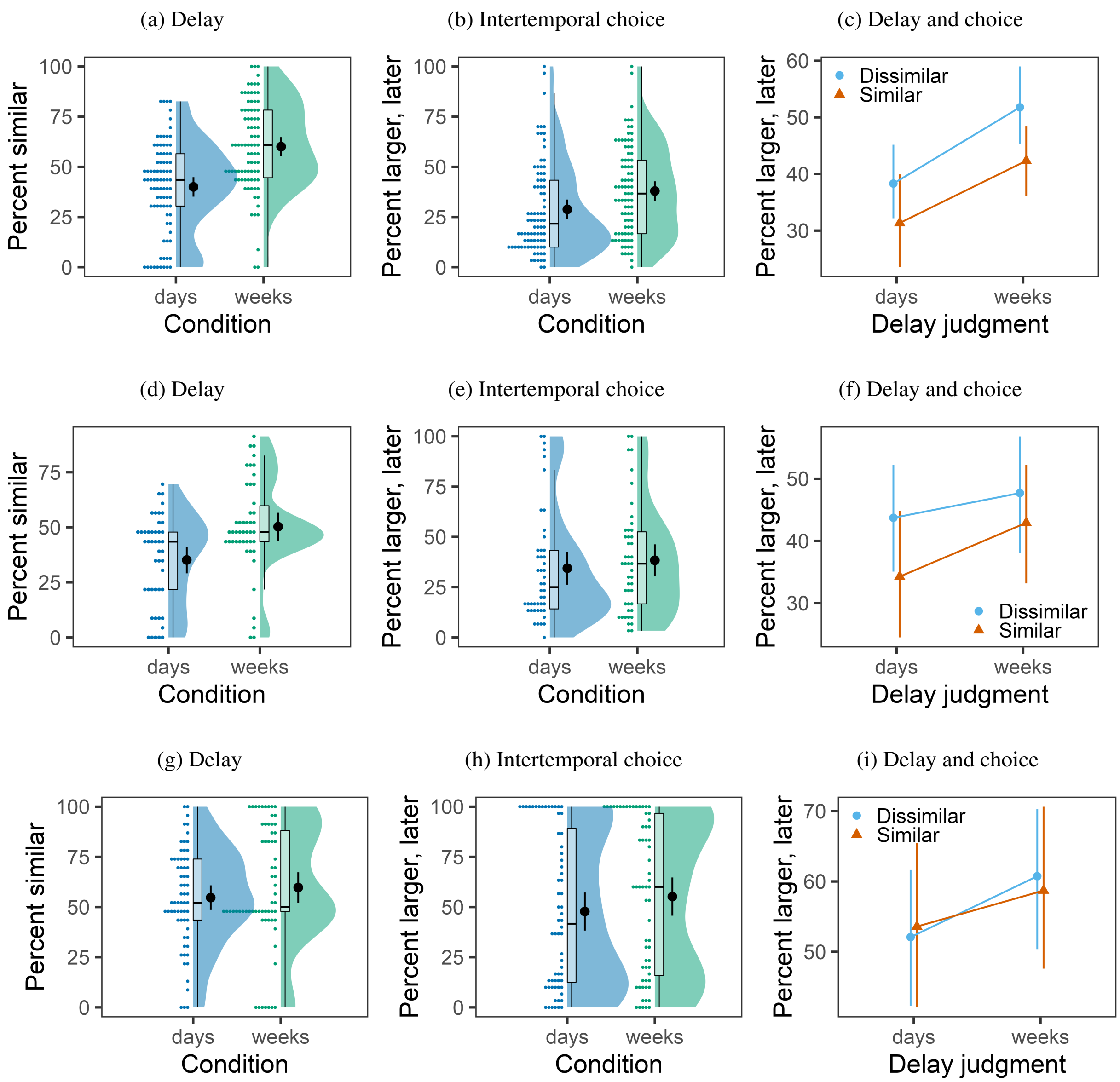

Figure 2. Delay framing similarity judgments, intertemporal choices, and interactions for days vs. weeks conditions for (a-c) data set 1 (N=86 per condition), (d-f) data set 2 (N=46 per condition), and (g-i) data set 3 (N=64 per condition). (a, d, g) Participants judged pairs of time delays as similar or dissimilar, and we calculated the mean percent similarity rating over all pairs for each participant. (b, e, h) Participants choose between smaller, sooner and larger, later options, and we calculated the mean percent choice for the larger, later per participant. (c, f, i) We calculated the percent choice for the larger, latter option conditional on whether participants judged the delay pair in the choice as similar or dissimilar for each condition. Small dots represent individual means for participants, filled circles/triangles and error bars represent condition means and 95% between-subjects confidence intervals, boxes represent condition interquartile ranges, horizontal lines represent condition medians, and whiskers represent 1.5 times the interquartile range.

was anecdotal evidence that numeracy and number line estimate errors were not correlated (r = .13, 95% CI [−.08, .32], $t(89)$ = 1.20, p = .235, BF_{10} = 0.47). Model selection procedures did not favor the interaction model with condition and numeracy or number line errors for either similarity judgments or intertemporal choice (Table S5). In data set 3, model selection procedures did not favor the interaction model with condition and numeracy for either similarity judgments or intertemporal choice (Table S5). Thus, numerical ability did not moderate the effect of condition on delay similarity judgments or intertemporal choices.

Summary

In this experiment, participants made delay similarity judgments and intertemporal choices with the time delays framed as either days or weeks. The results for the delay similarity experiment were mixed. In two of three data sets, participants in the weeks condition rated time delay pairs as more similar than participants in the days condition. This finding supports the notion that, though both frames have the same numerical ratio between time delays, the days condition results in a larger numerical difference, which resulted in those time delays being perceived as more dissimilar than weeks, even though the ratios and actual delays are identical. The MTurk data set showed no difference between the two framing conditions on similarity judgments. In one of three data sets, participants in the weeks condition also chose the larger, later option more frequently than those in the days condition. This follows from the delay similarity judgments because, if participants judge delays in the days condition as less similar than those in the weeks condition, they should be more likely to focus on time delays in their intertemporal choices, which would result in choosing the smaller, sooner option more frequently. Thus, the difference between conditions aligns with the similarity account of intertemporal choice for that data set. For the other two data sets, however, intertemporal choices did not differ between framing conditions, contradicting the similarity model account. Mediation analysis demonstrated that similarity judgments partially mediated the effect of framing condition on choice, but only in one of three data sets. Finally, numeracy did not moderate the effect of framing condition on similarity judgments or choice.

General Discussion

The present study investigated how the framing of reward amounts and time delays influenced people's similarity judgments and intertemporal choices. We found that framing reward amounts using larger units led participants to judge amounts as less similar and prefer the larger, later intertemporal choice option. Moreover, amount similarity judgments partially mediated the relationship between framing condition and intertemporal choice. In contrast, though framing time delays using larger units led participants to judge delays as less similar in two of three data sets, this led participants to prefer the larger, later intertemporal choice option in only one of the three data sets. Moreover, delay similarity judgments mediated the relationship between framing condition and intertemporal choice in only one data set. Finally, numeracy did not moderate the relationship between framing condition and similarity judgments or choice for both reward amounts and time delays.

Implications

The effect of unit framing on similarity judgments and intertemporal choices has a number of implications. First, we found that framing reward amounts and time delays using a scale with more individual units leads to fewer similarity judgments. This provides a potential explanation for the mechanism underlying the unit effect that leads people to perceive a greater difference between two values. The unit effect has been attributed to an increase in the salience of quantitative information while ignoring unit information for numbers (Pandelaere et al., 2011). That is, people only consider the size of the numbers and fail to consider that this quantitative information can be expressed in other units. Our findings provide a potential mechanism for the unit effect: similarity. Pairs of values that have the same numerical ratio but varying numerical differences produce different judgments of similarity. These similarity judgments then feed into decision processes for various forms of decision making such as intertemporal choice, risky choice, and strategic choice (Rubinstein, 1988; Leland, 2002, 2013).

Second, unit framing in intertemporal choices could potentially help people make better decisions. Impulsive decision making is a problem in many important decision-making domains such as diet, exercise, substance abuse, environmental sustainability, and saving for retirement (Hirsh et al., 2015; Knoll et al., 2015; Stevens, 2017). Framing reward amounts using scales that contain more units could nudge people into being less impulsive. Though framing monetary amounts as cents rather than dollars may not be practical in many contexts, there may be other situations in which using representations with more units could be useful in framing reward amounts. For instance, Camilleri and Larrick (2014) studied how describing vehicles on different scales influenced people's preferences for fuel-efficient vehicles that were more environmentally friendly. They found that describing the cost of gas for driving a fuel-efficient vehicle 100,000 miles led participants to prefer the fuel-efficient option compared to when a smaller scale of 15,000 miles was used. Framing amounts on a larger scale can thus lead people to prefer options that will promote environmental sustainability. Similarly, the manner in which time delays are framed can affect impulsivity. People wait for the larger, later option more often when the delay is communicated in terms of calendar dates compared to units of delay (e.g., days and weeks, Read et al., 2005; LeBoeuf, 2006) and when the delay is presented in units that contain a smaller scale (e.g., weeks vs. days, Siddiqui et al., 2018). Thus, stakeholder who may want to nudge people into choosing options that will provide them with greater benefits in the long term may do well to present delays in either date format or smaller-scaled units to decrease the perception of wait time.

Third, our results can extend to other domains of choice. The similarity approach has already been applied to risky choice (Rubinstein, 1988; Leland, 1994, 1998) and strategic choice in games (Leland, 2013), and it can easily extend to other forms of multi-attribute choice. For risky choice, framing options using a larger frame can affect subjective perception of risk. For instance, presenting the probability of disease risk using a 1-in-X format compared to an N-in-X×N format (e.g., 1 in 100 vs. 5 in 500) increases risk perception (Pighin et al., 2011; Oudhoff & Timmermans, 2015; Freeman et al., 2021). Our finding that framing reward amounts in larger units led to fewer similarity judgments between amount values suggests that presenting choice options using larger numerators and denominators (even though the ratios for options are equivalent) could lead people to consider the amount of risk for options to be less similar which would in turn lead them to choose the less risky option. In terms of consumer choice, our findings corroborate those that found that the use of expanded scales to describe product attributes can emphasize differences among similar products which can lead people to choose the option with the higher value on the scale (e.g., Pandelaere et al., 2011; Camilleri & Larrick, 2014; Cadario et al., 2016). Conversational logic may provide an explanation for this effect where a greater number of units translates to greater importance of the attribute under consideration because there is more detail about the attribute (Schwarz, 1994). Our finding that framing reward amounts in larger units led to fewer similarity judgments provides support for the notion that marketers would benefit from using larger scales to emphasize attribute differences when using comparative advertising because it would encourage consumers to judge product attributes to be less similar to one another.

Potential issues an limitations

Though the effects of unit framing on reward amounts were robust across data sets, the effects of unit framing on time delays were mixed. The first data set showed effects of framing on judgments and choice, the second data set only showed effects on judgments, and the third showed no effects. A possible explanation for the mixed results could be the lower ratio used for time delays compared to reward amounts. Specifically, we framed amount values in dollars and cents with a ratio of 1:100, whereas delay values were framed in days and weeks with a ratio of 1:7. The much smaller ratio for delays could have simply not been large enough to allow for the unit effect to be observed like for amounts. Research that found effects of unit framing on time delays used higher ratios of 1:24 for days vs. hours (Siddiqui et al., 2018). Thus, framing delays in units that would allow for larger ratios (e.g., hours vs. days or days vs. months) might result in stronger unit effects. Moreover, Siddiqui et al. (2018) reported that the unit framing effect only applied to time delays when the reward was hedonic (compared to utilitarian) in nature. Therefore, another reason why we did not find a preference for the larger, later option when we framed delays in larger compared to smaller units could have been due to participants considering the nature of our dollar amount rewards utilitarian instead of hedonic.

Additionally, the range of individual variation among participants was especially high in the third data set for the delay unit manipulation (Figure 2g-i), which could have led to the mixed findings for framing effects on similarity judgments and intertemporal choice. Individual variation in time perception can affect intertemporal choice (Takahashi, 2005; Kim & Zauberman, 2009; Wittmann, 2009; Zauberman et al., 2009). Thus, the participants in these data sets could just be an anomalous sample. However, there are two notable features of this sample that could also account for the difference. First, the third data set was drawn from a pool of Mechanical Turk workers rather than university students. Differences that we observed in this data set could have resulted from a lack of generalizability from students to the general public (McNemar, 1946; Gordon et al., 1986) or from lower quality data from Mechanical Turk workers (Aruguete et al., 2019; Chmielewski & Kucker, 2020; Gupta et al., 2021). Second, this data set was collected during the COVID-19 pandemic of 2021, which can influence people's decision making, including risky choice (Yue et al., 2020). Thus, this data set may not be indicative of non-pandemic decision making.

We predicted that numeracy would moderate the relationship between framing condition and similarity judgments and intertemporal choice based on work that found more numerate individuals were less susceptible to framing effects (Peters, 2012; Ghazal et al., 2014). However, we did not find these effects. One reason for this discrepancy could be because we used the multiple choice version of the Berlin Numeracy Test to assess participants' numeracy, which may not have had sufficient sensitivity to detect numeracy effects. Future research should explore if other numeracy measures yield findings similar to those of the present study.

Finally, participants made similarity judgments for amount and delay values using binary response options (i.e., similar or dissimilar). A drawback of this approach is that it does not capture instances in which participants may consider the values for both amount and delay attributes as similar but with one attribute *more* similar than the other. When both attributes are considered similar, similarity models suggest that people will either make a choice at random (Leland, 2002) or rely on other choice criteria (Rubinstein, 2003). Using a continuous similarity measure that ranges from similar on one end to dissimilar on the other would allow us to pinpoint the degree of similarity for attributes, resulting in the attribute with a lower degree of similarity being considered the dissimilar attribute. This may then allow similarity models to better predict intertemporal choices. Moreover, for future studies, continuous measures of similarity may assess the degree of similarity for attributes more accurately and sensitively, leading to a better ability to detect potential unit framing effects.

Conclusions

Intertemporal choices are ubiquitous in life. The present research investigated how using different unit frames for reward amounts and time delays affected people's similarity judgments for option attributes and subsequent intertemporal choice. We showed that framing reward amounts in larger units caused people to judge amounts as less similar and prefer the larger, later option but that this effect was mixed when it came to time delays. Taken together, our findings suggest that unit choice matters in the communication of intertemporal choice options and that similarity judgments play an underlying role in people's choices. While certain individual differences could potentially attenuate the effect of unit frames, numeracy may be an insufficient buffer against it.

Acknowledgments

This research was funded by National Science Foundation grants (NSF-1062045, NSF-1658837). We would like to thank Yesenia Delgadillo-Orozco, Maddie Hemenway, Stephanie Marsh, Carly Moran, Nina Radulovic, Alex Ruhlman, Alexis Saltzman, Sam Wenzl, Austin Williams, and Molly Zuercher for testing participants and Billy Lim and London Wolff for comments on an early draft.

References

Amlung, M., Vedelago, L., Acker, J., Balodis, I., & MacKillop, J. (2017). Steep delay discounting and addictive behavior: A meta-analysis of continuous associations. *Addiction*, *112*(1), 51–62. https://doi.org/10.1111/add.13535

Andraszewicz, S., Scheibehenne, B., Rieskamp, J., Grasman, R., Verhagen, J., & Wagenmakers, E.-J. (2015). An introduction to Bayesian hypothesis testing for management research. *Journal of Management*, *41*(2), 521–543. https://doi.org/10.1177/0149206314560412

Aruguete, M. S., Huynh, H., Browne, B. L., Jurs, B., Flint, E., & McCutcheon, L. E. (2019). How serious is the 'carelessness' problem on Mechanical Turk? *International Journal of Social Research Methodology*, *22*(5), 441–449. https://doi.org/10.1080/13645579.2018.1563966

Atlas, S. A., & Bartels, D. M. (2018). Periodic pricing and perceived contract benefits. *Journal of Consumer Research*, *45*(2), 350–364. https://doi.org/10.1093/jcr/ucx127

Aust, F., & Barth, M. (2017). *papaja: Create APA manuscripts with R Markdown*. https://github.com/crsh/papaja

Bagchi, R., & Davis, D. F. (2016). The role of numerosity in judgments and decision-making. *Current Opinion in Psychology*, *10*, 89–93. https://doi.org/10.1016/j.copsyc.2015.12.010

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bickel, W. K., Athamneh, L. N., Basso, J. C., Mellis, A. M., DeHart, W. B., Craft, W. H., & Pope, D. (2019). Excessive discounting of delayed reinforcers as a trans-disease process. *Current Opinion in Psychology*, *30*, 59–64. https://doi.org/10.1016/j.copsyc.2019.01.005

Burson, K. A., Larrick, R. P., & Lynch, J. G. (2009). Six of one, half dozen of the other: Expanding and contracting numerical dimensions produces preference reversals. *Psychological Science*, *20*(9), 1074–1078. https://doi.org/10.1111/j.1467-9280.2009.02394.x

Buschena, D. E., & Zilberman, D. (1999). Testing the effects of similarity on risky choice: Implications for violations of expected utility. *Theory and Decision*, *46*(3), 251–276. https://doi.org/10.1023/A:1005066504527

Cadario, R., Parguel, B., & Benoit-Moreau, F. (2016). Is bigger always better? The unit effect in carbon emissions information. *International Journal of Research in Marketing*, *33*(1), 204–207. https://doi.org/10.1016/j.ijresmar.2015.10.002

Camilleri, A. R., & Larrick, R. P. (2014). Metric and scale design as choice architecture tools. *Journal of Public Policy & Marketing*, *33*(1), 108–125. https://doi.org/10.1509/jppm.12.151

Chen, S.-F. S., Monroe, K. B., & Lou, Y.-C. (1998). The effects of framing price promotion messages on consumers' perceptions and purchase intentions. *Journal of Retailing*, *74*(3), 353–372. https://doi.org/10.1016/S0022-4359(99)80100-6

Chmielewski, M., & Kucker, S. C. (2020). An MTurk crisis? Shifts in data quality and the impact on study results. *Social Psychological and Personality Science*, *11*(4), 464–473. https://doi.org/10.1177/1948550619875149

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin Numeracy Test. *Judgment and Decision Making*, *7*(1), 25–47.

Coulter, K. S., Choi, P., & Monroe, K. B. (2012). Comma N' cents in pricing: The effects of auditory representation encoding on price magnitude perceptions. *Journal of Consumer Psychology*, *22*(3), 395–407. https://doi.org/10.1016/j.jcps.2011.11.005

DeHart, W. B., & Odum, A. L. (2015). The effects of the framing of time on delay discounting. *Journal of the Experimental Analysis of Behavior*, *103*(1), 10–21. https://doi.org/10.1002/jeab.125

Delgado-Gomez, D., Baca-Garcia, E., Aguado, D., Courtet, P., & Lopez-Castroman, J. (2016). Computerized Adaptive Test vs. Decision trees: Development of a support decision system to identify suicidal behavior. *Journal of Affective Disorders*, *206*, 204–209. https://doi.org/10.1016/j.jad.2016.07.032

Doyle, J. R. (2013). Survey of time preference, delay discounting models. *Judgment and Decision Making*, *8*(2), 116–135.

Freeman, A. L. J., Kerr, J., Recchia, G., Schneider, C. R., Lawrence, A. C. E., Finikarides, L., Luoni, G., Dryhurst, S., & Spiegelhalter, D. (2021). Communicating personalized risks from COVID-19: Guidelines from an empirical study. *Royal Society Open Science*, *8*(4). https://doi.org/10.1098/rsos.201721

Ghazal, S., Cokely, E. T., & Garcia-Retamero, R. (2014). Predicting biases in very highly educated samples: Numeracy and metacognition. *Judgment and Decision Making*, *9*(1), 15–34.

Goh, F. W., & Stevens, J. R. (2021). Attribute-based choice. In R. Viale (Ed.), *Routledge Handbook of Bounded Rationality* (pp. 242–253). Routledge.

Goldstein, D. G., Hershfield, H. E., & Benartzi, S. (2016). The illusion of wealth and its reversal. *Journal of Marketing Research*, *53*(5), 804–813. https://doi.org/10.1509/jmr.14.0652

Gordon, M. E., Slade, L. A., & Schmitt, N. (1986). The "science of the sophomore" revisited: From conjecture to empiricism. *Academy of Management Review*, *11*(1), 191–207. https://doi.org/10.5465/AMR.1986.4282666

Gourville, J. T. (2003). The effects of monetary magnitude and level of aggregation on the temporal framing of price. *Marketing Letters*, *14*(2), 125–135. https://doi.org/10.1023/A:1025467002310

Gupta, N., Rigotti, L., & Wilson, A. (2021). The experimenters' dilemma: Inferential preferences over populations. *arXiv*, 2107.05064. http://arxiv.org/abs/2107.05064

Hardisty, D. J., & Weber, E. U. (2009). Discounting future green: Money versus the environment. *Journal of Experimental Psychology: General*, *138*(3), 329–340. https://doi.org/10.1037/a0016433

Hershfield, H. E., Shu, S., & Benartzi, S. (2020). Temporal reframing and participation in a savings program: A field experiment. *Marketing Science*, *39*(6), 1039–1051. https://doi.org/10.1287/mksc.2019.1177

Hirsh, J. L., Costello, M. S., & Fuqua, R. W. (2015). Analysis of delay discounting as a psychological measure of sustainable behavior. *Behavior and Social Issues*, *24*(1), 187–202. https://doi.org/10.5210/bsi.v24i0.5906

Imai, K., Keele, L., & Tingley, D. (2010a). A general approach to causal mediation analysis. *Psychological Methods*, *15*(4), 309–334. http://imai.princeton.edu/research/BaronKenny.html

Imai, K., Keele, L., Tingley, D., & Yamamoto, T. (2011). Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies. *American Political Science Review*, *105*(4), 765–789. http://imai.princeton.edu/research/mediationP.html

Imai, K., Keele, L., & Yamamoto, T. (2010b). Identification, inference, and sensitivity analysis for causal mediation effects. *Statistical Science*, *25*(1), 51–71. http://imai.princeton.edu/research/mediation.html

Imai, K., & Yamamoto, T. (2013). Identification and sensitivity analysis for multiple causal mechanisms: Revisiting evidence from framing experiments. *Political Analysis*, *21*(2), 141–171. http://imai.princeton.edu/research/medsens.html

Kay, M. (2021). *ggdist: Visualizations of distributions and uncertainty*. https://doi.org/10.5281/zenodo.3879620

Kim, B. K., & Zauberman, G. (2009). Perception of anticipatory time in temporal discounting. *Journal of Neuroscience, Psychology, and Economics*, *2*(2), 91–101. https://doi.org/10.1037/a0017686

Kim, K., & McKinnon, L. M. (2020). Framing financial advertising: Message effectiveness in intertemporal choice. *Journal of Marketing Communications*, *26*(3), 328–342. https://doi.org/10.1080/13527266.2018.1476400

Knoll, M. A. Z., Appelt, K. C., Johnson, E. J., & Westfall, J. E. (2015). Time to retire: Why Americans claim benefits early & how to encourage delay. *Behavioral Science & Policy*, *1*(1), 53–62. https://doi.org/10.1353/bsp.2015.0003

Kwong, J. Y. Y., & Wong, K. F. E. (2006). The role of ratio differences in the framing of numerical information. *International Journal of Research in Marketing*, *23*(4), 385–394. https://doi.org/10.1016/j.ijresmar.2006.07.001

LeBoeuf, R. A. (2006). Discount rates for time versus dates: The sensitivity of discounting to time-interval description. *Journal of Marketing Research*, *43*(1), 59–72.

Leland, J. W. (2013). Equilibrium selection, similarity judgments, and the "nothing to gain/nothing to lose" effect. *Journal of Behavioral Decision Making*, *26*(5), 418–428. https://doi.org/10.1002/bdm.1772

Leland, J. W. (1994). Generalized similarity judgments: An alternative explanation for choice anomalies. *Journal of Risk and Uncertainty*, *9*(2), 151–172.

Leland, J. W. (1998). Similarity judgments in choice under uncertainty: A reinterpretation of the predictions of regret theory. *Management Science*, *44*(5), 659–672. https://doi.org/10.1287/mnsc.44.5.659

Leland, J. W. (2002). Similarity judgments and anomalies in intertemporal choice. *Economic Inquiry*, *40*(4), 574–581. https://doi.org/10.1093/ei/40.4.574

Luan, S., Schooler, L. J., & Gigerenzer, G. (2011). A signal-detection analysis of fast-and-frugal trees. *Psychological Review*, *118*(2), 316–338. https://doi.org/10.1037/a0022684

Lüdecke, D., Makowski, D., Waggoner, P., & Patil, I. (2020). performance: Assessment of regression models performance. *CRAN*. https://doi.org/10.5281/zenodo.3952174

Makowski, D., Ben-Shachar, M. S., & Lüdecke, D. (2019). bayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. *Journal of Open Source Software*, *4*(40), 1541. https://doi.org/10.21105/joss.01541

McNemar, Q. (1946). Opinion-attitude methodology. *Psychological Bulletin*, *43*(4), 289–374. https://doi.org/10.1037/h0060985

Morey, R. D., & Rouder, J. N. (2015). *BayesFactor: Computation of Bayes factors for common designs*. https://CRAN.R-project.org/package=BayesFactor

Morris, K. L., & Perna, F. M. (2018). Decision tree model vs traditional measures to identify patterns of sun-protective behaviors and sun sensitivity associated with sunburn. *JAMA Dermatology*, *154*(8), 897. https://doi.org/10.1001/jamadermatol.2018.1646

Müller, K. (2017). *here: A simpler way to find your files*. https://CRAN.R-project.org/package=here

Oudhoff, J. P., & Timmermans, D. R. M. (2015). The effect of different graphical and numerical likelihood formats on perception of likelihood and choice. *Medical Decision Making*, *35*(4), 487–500. https://doi.org/10.1177/0272989X15576487

Pacini, R., & Epstein, S. (1999). The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon. *Journal of Personality and Social Psychology*, *76*(6), 972–987. https://doi.org/10.1037/0022-3514.76.6.972

Pandelaere, M., Briers, B., & Lembregts, C. (2011). How to make a 29% increase look bigger: The unit effect in option comparisons. *Journal of Consumer Research*, *38*(2), 308–322. https://doi.org/10.1086/659000

Park, I., & Cho, S. (2019). The influence of number line estimation precision and numeracy on risky financial decision making. *International Journal of Psychology*, *54*(4), 530–538. https://doi.org/10.1002/ijop.12475

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Pelham, B. W., Sumarta, T. T., & Myaskovsky, L. (1994). The easy path from many to much: The numerosity heuristic. *Cognitive Psychology*, *26*(2), 103–133. https://doi.org/10.1006/cogp.1994.1004

Peters, E. (2012). Beyond comprehension: The role of numeracy in judgments and decisions. *Current Directions in Psychological Science*, *21*(1), 31–35. https://doi.org/10.1177/0963721411429960

Peters, E., Slovic, P., Västfjäll, D., & Mertz, C. K. (2008). Intuitive numbers guide decisions. *Judgment and Decision Making*, *3*, 619–635.

Pighin, S., Savadori, L., Barilli, E., Cremonesi, L., Ferrari, M., & Bonnefon, J.-F. (2011). The 1-in-X effect on the subjective assessment of medical probabilities. *Medical Decision Making*, *31*(5), 721–729. https://doi.org/10.1177/0272989X11403490

R Core Team. (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rahimi-Golkhandan, S., Garavito, D. M. N., Reyna-Brainerd, B. B., & Reyna, V. F. (2017). A Fuzzy-Trace Theory of risk and time preferences in decision making: Integrating cognition and motivation. In J. R. Stevens (Ed.), *Impulsivity: How time and risk influence decision making, Nebraska Symposium on Motivation* (Vol. 64, pp. 115–144). Springer International Publishing. https://doi.org/10.1007/978-3-319-51721-6_4

Read, D., Frederick, S., Orsel, B., & Rahman, J. (2005). Four score and seven years from now: The date/delay effect in temporal discounting. *Management Science*, *51*(9), 1326–1335. https://doi.org/10.1287/mnsc.1050.0412

Regenwetter, M., Cavagnaro, D. R., Popova, A., Guo, Y., Zwilling, C., Lim, S. H., & Stevens, J. R. (2018). Heterogeneity and parsimony in intertemporal choice. *Decision*, *5*(2), 63–94. https://doi.org/10.1037/dec0000069

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225–237. https://doi.org/10.3758/PBR.16.2.225

Rubinstein, A. (1988). Similarity and decision-making under risk (Is there a utility theory resolution to the Allais paradox?). *Journal of Economic Theory*, *46*(1), 145–153. https://doi.org/10.1016/0022-0531(88)90154-8

Rubinstein, A. (2003). "Economics and psychology"? The case of hyperbolic discounting. *International Economic Review*, *44*(4), 1207–1216. https://doi.org/10.1111/1468-2354.t01-1-00106

Schley, D. R., Lembregts, C., & Peters, E. (2017). The role of evaluation mode on the unit effect. *Journal of Consumer Psychology*, *27*(2), 278–286. https://doi.org/10.1016/j.jcps.2016.07.001

Schönbrodt, F. D., Wagenmakers, E.-J., Zehetleitner, M., & Perugini, M. (2017). Sequential hypothesis testing with Bayes factors: Efficiently testing mean differences. *Psychological Methods*, *22*(2), 322–339. https://doi.org/10.1037/met0000061

Schwarz, N. (1994). Judgment in a social context: Biases, shortcomings, and the logic of conversation. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 26, pp. 123–162). Academic Press. https://doi.org/10.1016/S0065-2601(08)60153-7

Shen, L., & Urminsky, O. (2013). Making sense of nonsense: The visual salience of units determines sensitivity to magnitude. *Psychological Science*, *24*(3), 297–304. https://doi.org/10.1177/0956797612451470

Siddiqui, R. A., Monga, A., & Buechel, E. C. (2018). When intertemporal rewards are hedonic, larger units of wait time boost patience. *Journal of Consumer Psychology*, *28*(4), 612–628. https://doi.org/10.1002/jcpy.1019

Skylark, W., Farmer, G., & Bahemia, N. (2021). Inference and preference in intertemporal choice. *Judgment and Decision Making*, *16*(2), 422–459. https://doi.org/10.17863/CAM.66563

Stevens, J. R. (2016). Intertemporal similarity: Discounting as a last resort. *Journal of Behavioral Decision Making*, *29*(1), 12–24. https://doi.org/10.1002/bdm.1870

Stevens, J. R. (2017). *Impulsivity: How Time and Risk Influence Decision Making*. Springer.

Stevens, J. R., & Soh, L.-K. (2018). Predicting similarity judgments in intertemporal choice with machine learning. *Psychonomic Bulletin & Review*, *25*(2), 627–635. https://doi.org/10.3758/s13423-017-1398-1

Story, G., Vlaev, I., Seymour, B., Darzi, A., & Dolan, R. (2014). Does temporal discounting explain unhealthy behavior? A systematic review and reinforcement learning perspective. *Frontiers in Behavioral Neuroscience*, *8*, 76. https://doi.org/10.3389/fnbeh.2014.00076

Takahashi, T. (2005). Loss of self-control in intertemporal choice may be attributable to logarithmic time-perception. *Medical Hypotheses*, *65*(4), 691–693. https://doi.org/10.1016/j.mehy.2005.04.040

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). mediation: R package for causal mediation analysis. *Journal of Statistical Software*, *59*(5), 1–38. http://www.jstatsoft.org/v59/i05/

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*, *14*(5), 779–804. https://doi.org/10.3758/BF03194105

Weber, E. U., Johnson, E. J., Milch, K. F., Chang, H., Brodscholl, J. C., & Goldstein, D. G. (2007). Asymmetric discounting in intertemporal choice: A query-theory account. *Psychological Science*, *18*(6), 516–523.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wittmann, M. (2009). The inner experience of time. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1525), 1955–1967. https://doi.org/10.1098/rstb.2009.0003

Yamagishi, K. (1997). When a 12.86% mortality is more dangerous than 24.14%: Implications for risk communication. *Applied Cognitive Psychology*, *11*(6), 495–506. https://doi.org/10.1002/(SICI)1099-0720(199712)11:6%3C495::AID-ACP481%3E3.0.CO;2-J

Yue, P., Gizem Korkmaz, A., & Zhou, H. (2020). Household financial decision making amidst the COVID-19 pandemic. *Emerging Markets Finance and Trade*, *56*(10), 2363–2377. https://doi.org/10.1080/1540496X.2020.1784717

Zauberman, G., Kim, B. K., Malkoc, S. A., & Bettman, J. R. (2009). Discounting time and time discounting: Subjective time perception and intertemporal preferences. *Journal of Marketing Research*, *46*(4), 543–556.