Journal of Experimental Psychology: General

Two Modes of Social Impressions and Their Effects on Choice

Leor M. Hackel and Peter Mende-Siedlecki

Online First Publication, May 18, 2023. https://dx.doi.org/10.1037/xge0001432

CITATION
Hackel, L. M., & Mende-Siedlecki, P. (2023, May 18). Two Modes of Social Impressions and Their Effects on Choice. *Journal of Experimental Psychology: General*. Advance online publication. https://dx.doi.org/10.1037/xge0001432

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General

© 2023 American Psychological Association
ISSN: 0096-3445

https://doi.org/10.1037/xge0001432

Two Modes of Social Impressions and Their Effects on Choice

Leor M. Hackel[1] and Peter Mende-Siedlecki[2]

[1] Department of Psychology, University of Southern California

[2] Department of Psychological and Brain Sciences, University of Delaware

Our memories of other people shape how we interact with them. Yet, even when we forget exactly what others said or did, we often remember impressions that capture a general gist of their behavior—whether they were forthright, friendly, or funny. Drawing on fuzzy trace theory, we propose two modes of social impression formation: impressions formed based on ordinal gist ("more competent," "less competent") or categorical gist ("competent," "incompetent"). In turn, we propose that people gravitate toward the simplest representation available and that different modes of memory have distinct consequences for social decisions. Specifically, ordinal impressions lead people to make decisions based on an individual's standing relative to others, whereas categorical impressions lead people to make decisions based on discrete classifications that interpret behavior. In four experiments, participants learned about two groups of individuals who differed in their competence (Studies 1a, 2, and 3) or generosity (Study 1b). When participants encoded impressions as ordinal rankings, they preferred to hire or help a relatively good target from a low-performing group over a relatively bad target from a high-performing group, even though both targets behaved identically and accuracy was incentivized. However, when participants could use categorical boundaries to interpret behavior, this preference was eliminated. In a final experiment, changing the category participants used to encode others' generosity changed their impressions, even when accounting for memory for verbatim details. This work links social impressions to theories of mental representation in memory and judgment, highlighting how distinct representations support divergent patterns of social decision-making.

Public Significance Statement

People choose who to hire or help based on fuzzy impressions of other people—for instance, hiring individuals they see as competent and helping those they see as generous. This study demonstrates that people can form these impressions in two distinct ways: people may either rank how individuals compare relatively to one another (e.g., a student who ranks first in a class vs. one who ranks third in a class) or assign individuals to descriptive categories that help make sense of their behavior (e.g., two "A" grades). These different routes can lead people to make very different decisions about who to help or hire, even when decision-makers are exposed to the same information and when they recall the same details of others' behavior.

Keywords: impression formation, social cognition, memory, decision-making, fuzzy trace theory

Supplemental materials: https://doi.org/10.1037/xge0001432.supp

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Our memories of other people shape how we interact with them, but we can remember others' behavior at multiple levels of abstraction. If a student performs well in a course, their instructor could remember the specific scores the student received; the student's ranking at the top of the class; or simply that the student was an "A" student. These memories all reflect the student's competence, but with different levels of detail that might lead to different decisions. For instance, in another semester, a second student might receive the same scores while ranking fifth in the class, due to the presence of even better-performing students. If the instructor recalls

Leor M. Hackel https://orcid.org/0000-0002-7959-0642

Results from these studies were presented at the Social & Affective Neuroscience Society Annual Meeting in 2022 and the Association for Psychological Science Annual Meeting in 2022. We thank David Kalkstein for insightful comments on the article.

All studies were preregistered, including sample size, measures, exclusion criteria, and analysis plan. Preregistration documents are available at https://aspredicted.org/4xh55.pdf, https://aspredicted.org/LJZ_2KL, https://aspredicted.org/7N8_GG6, and https://aspredicted.org/FXG_663. We report all measures, manipulations, and data exclusion criteria. Deidentified data and analysis code are available at https://osf.io/xwvue/?view_only=250a1baab4944912929afbdccc314880.

Leor M. Hackel and Peter Mende-Siedlecki designed the studies. Leor M. Hackel served as lead for data curation, formal analysis, visualization, and writing–original draft. Peter Mende-Siedlecki served in a supporting role for writing–original draft. Leor M. Hackel and Peter Mende-Siedlecki contributed equally to conceptualization, writing–review and editing, and methodology.

Correspondence concerning this article should be addressed to Leor M. Hackel, Department of Psychology, University of Southern California, 3620 South McClintock Avenue, Los Angeles, CA 90089, United States. Email: lhackel@usc.edu

student performance in terms of rankings, they might recommend the first student for a scholarship over the second, even if both behaved identically. To what extent does the level of detail with which people encode impressions shape their social decisions?

To answer this question, we draw on fuzzy trace theory (Brainerd & Reyna, 2001; Reyna & Brainerd, 1995)—a theory of memory, judgment, and decision-making—to propose that people encode social impressions in memory in two modes: ordinal impressions, which reflect a person's traits relative to others, and categorical impressions, which reflect a person's traits in terms of discrete categories. In turn, we argue that these modes of impression formation have distinct consequences for social decisions. For instance, ordinal impressions would lead people to prefer an individual who stood out positively against others over an individual who stood out negatively against others, even if these two individuals showed equivalent behavior. In contrast, categorical impressions can avoid this tendency by assigning both individuals to the same category.

Ordinal and Categorical Gist in Decision-Making

Fuzzy trace theory posits that people simultaneously form two types of memories of events in parallel: "verbatim" representations, which capture precise details of an event, and "gist" representations, which capture the essential meaning of an event (Brainerd & Reyna, 2001; Reyna & Brainerd, 1995). Whereas verbatim representations are literal and concrete, gists are abstract and convey broader meaning. In this framework, verbatim and gist representations exist on a spectrum, with two types of gist representations filling out this spectrum. "Ordinal gist" moves from verbatim representations of information (e.g., "a 50% chance of rain" vs. "a 30% chance of rain") to ordinal rankings (e.g., "more likely" vs. "less likely"). This spectrum then expands to include "categorical gist," which involves even simpler representations of categories (e.g., "some chance" vs. "no chance"; Corbin et al., 2015; Mills et al., 2008; Reyna et al., 2014).[1] Different tasks can promote different representations, as people typically gravitate to the simplest representation of a task or problem they can use (Corbin et al., 2015). Although people process all three types of information and use all of them in decision-making, people tend to give greater weight to the simplest gist available when making choices (Reyna et al., 2021).

Crucially, different gist representations can give rise to different decisions. For instance, teens who think about risk categorically ("no risk is better than some risk") are more likely to avoid risk than teens who think about risk ordinally ("less risk is better than more risk"; Mills et al., 2008). Accordingly, changing the gists people use to encode information can alter the decisions they make. For instance, classic framing effects in decision-making can be enhanced or eliminated by changing gist representations (Reyna et al., 2014). In the "dread disease" problem (Kahneman & Tversky, 1984), people typically avoid risk with a gain frame ("200 people can be saved for sure, or there will be a one-third probability 600 people will be saved and a two-third probability no one will be saved"), but people typically seek risk with a loss frame ("400 people will die for sure, or there will be a two-third probability 600 people die and a one-third probability no one dies"). According to fuzzy trace theory, this is because the simplest gist representation of the gain frame is "some are saved" versus "either some or none are saved," while the simplest representation of the loss frame is "some people die" versus "either some die or none die." People prefer "some are saved" to "none are saved" but "none die" to "some die," leading to framing effects. Therefore, simply removing redundant information that evokes the category "none" (i.e., simply stating "there will be a two-third probability 600 people die") eliminates the framing effect (Reyna et al., 2014). Altogether, decision-making depends on gist: the type of representation that people use when encoding, storing, or retrieving information shapes the decisions they make based on that information.

Social Impressions as Gist Representations

We propose that social impressions reflect gists of social behavior, and that social decisions accordingly depend on the type of gist representation encoded. Like other gists in fuzzy trace theory, trait impressions such as "generous" or "competent" describe the essential meaning of behavior rather than its episodic details (Fiske, 1993). In particular, traits have been theorized to reflect goal-based categories, such that traits describe the social meaning of many concrete acts bound by an abstract goal (Read et al., 1990). For instance, a person might be labeled as "trustworthy" for repaying a loan, keeping a secret, or returning a lost object. These acts differ in their particulars but reflect a similar social significance.

Consistent with fuzzy trace theory, people form both verbatim and gist memories of social behavior. For instance, when people read about others' actions, they encode in memory both the exact details of a sentence ("The barista tripped and spilled the drinks") and trait concepts that summarize the episode ("clumsy"; Todorov & Uleman, 2002, 2003; Winter & Uleman, 1984). Fuzzy trace theory further predicts that gist memories are longer-lasting and more easily accessed than verbatim memories (Brainerd & Reyna, 2001; Corbin et al., 2015). Consistent with this view, people can form and remember hundreds of positive or negative impressions of others (Falvello et al., 2015) though they typically forget the exact behaviors that gave rise to those impressions (Todorov & Uleman, 2002).

Accordingly, we propose that social impressions can reflect ordinal or categorical gist. Supporting the prediction that impression formation may depend on ordinal encoding, trait impressions are often relative: whether a student is considered "talkative" might depend on how often the rest of a class participates, and whether a donor is considered "generous" might depend on how much money other fundraiser attendees have given. People may thus encode traits in terms of ordinal rankings relative to a reference point ("more generous" vs. "less generous"). Indeed, prior work on social impression formation supports this proposition (Dotsch et al., 2017; Mende-Siedlecki, Baron, & Todorov, 2013; Sanfey, 2009).

However, in other cases, people may use categories to encode behaviors. For instance, a workplace might have reference standards for what counts as "unacceptable," "acceptable," or "excellent" performance or as "early," "on time," or "late" arrival. Alternatively, an investor might be evaluated based on whether they gained or lost money—a categorical distinction that can lead to impressions of competence or incompetence. In these cases, trait impressions may reflect judgments relative to a categorical reference standard, rather than judgments reflecting a ranking relative to salient others. According to fuzzy trace theory, people would gravitate toward

[1] Although "some" versus "none" contain ordinal information in a mathematical sense, categorical gist in fuzzy trace theory assigns stimuli to discrete categories whereas ordinal gist maintains the ordinal rankings of all stimuli.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

this simpler categorical representation when possible. Past work suggests that people use social categories to form impressions of group members (e.g., Fiske & Neuberg, 1990), but less is known about the consequences of encoding a trait impression itself in terms of category boundaries.

Consequences for Social Decisions

How might these two modes of impression formation shape social decisions? Consider an employer evaluating two candidates for a job on separate days, both of whom are equally competent. One candidate happens to be evaluated along with a set of unimpressive applications, while the other candidate happens to be evaluated along with a set of impressive applications. If people form impressions as ordinal gists, they would judge the "relatively competent" candidate positively and the "relatively incompetent" candidate negatively. In turn, when deciding which of these candidates to hire, employers may fail to rescale their impressions to account for the original differences in contexts. As a result, they would hire the "relatively competent" candidate over the "relatively incompetent" candidate. However, this preference would not emerge if the employer forms impressions based on discrete categories they apply consistently to both candidates. For instance, if the employer applies categorical reference standards to each applicant in a consistent manner (e.g., "noncompetitive," "borderline," "competitive"), then gist representations of these two individuals would be identical. Hiring decisions between these individuals should therefore be less likely to reflect the learning context.

Recent research supports the prediction that people fail to rescale context-dependent memories when comparing ordinal rankings learned in different contexts. For instance, Sharif and Oppenheimer (2016) showed participants a toy race car driving at moderate speed. In addition, participants saw either a faster or slower race car. Participants later viewed a supposedly new car, which was in fact identical to the initial moderate-speed car (a "decoy"). When the moderate car had been encoded as relatively slow, participants ranked it as slower than the decoy car, and when the moderate car had been relatively fast, participants ranked it as faster than the decoy car. Thus, memories appear to retain their original ordinal encoding in later judgments, with the power to shape subsequent choices between targets encountered under *different* frames of reference.

Context-dependence exerts a similar influence on decision-making following reward-based learning, such that rewards are encoded as "good" or "bad" relative to contextual baselines (Palminteri et al., 2015; Palminteri & Lebreton, 2021). For instance, Palminteri et al. (2015) asked participants to learn the reward value of virtual "slot machines." In a "gain" context, a "good" slot offered a high chance of a $1 gain and a "bad" slot offered a low chance of a $1 gain. In a "loss" context, a "bad" slot machine offered a high chance of a $1 loss and a "good" slot offered a low chance of a $1 loss. In later choices made without feedback, participants preferred the "good-loss" slot over the "bad-gain" slot, even though the "good-loss" slot was objectively worse. Participants thus encoded options relative to reference points and did not rescale representations in later choices.

Together, these lines of work support the idea that impressions formed as ordinal gists may shape social decision-making. Indeed, Sharif and Oppenheimer similarly found that participants used context-dependent memories when judging a singer's audition or a job applicant's resume (Sharif & Oppenheimer, 2016, 2021), suggesting that this reference-dependence effect extends to social contexts. However, if people gravitate to the simplest representation available in a task, they may make different choices when they can apply categorical reference standards to another's behavior—for instance, avoiding the reference-dependent preferences described above.

Overview of Studies

Altogether, fuzzy trace theory provides a foundation for (a) understanding why people form social impressions (impressions are gist memories that turn rich details into simple meaning); (b) characterizing the nature of mental representation underlying those impressions (a spectrum of abstraction including verbatim details, ordinal scales, or discrete categories); (c) predicting which representations people use in different settings (people gravitate toward the simplest representations afforded by a task); and (d) predicting how these representations impact choice (changing a perceiver's gist can change their decisions, even when the underlying information remains the same). Although fuzzy trace theory has been used to explain nonsocial decisions, it thus offers a useful framework for understanding impression formation and predicting social decisions.

We generated the following predictions, drawing on fuzzy trace theory. First, when people cannot sort others' behavior into discrete categories, their decisions will reflect ordinal rankings of others' behavior (e.g., "donated more" vs. "donated less"); second, when people have categories available to describe others' behavior, they will invoke this simpler representation and their decisions will reflect categorical gist (e.g., "lost money" vs. "gained money"); third, when different people use different categorical gists to make sense of the same behavior, their impressions and choices will differ, even if ordinal rankings and verbatim details remain the same.

In Studies 1a and 1b, we tested our first prediction, asking whether ordinal gist shapes decision-making when people lack reference standards they can use to categorize behavior. Participants learned about two individuals who were equally competent (Study 1a) or equally generous (Study 1b) in an economic game; however, these individuals appeared among others who made them look relatively good or relatively bad. In later choices to hire (Study 1a) or share with (Study 1b) these individuals, we examined whether participants would prefer the *relatively good* individual over the *relatively bad* individual. By testing this question across the domains of competence and generosity, we examined two key dimensions of social perception that shape social decision-making (e.g., Fiske et al., 2007).

Next, Studies 2–3 tested our second prediction: We asked whether this reference-dependent preference would be eradicated when people encoded impressions in a categorical, rather than ordinal, manner. In Study 2, participants were randomly assigned to either an "ordinal" condition, in which they had no reference standards with which to interpret behavior in a novel task, or a "categorical" condition, in which they were given category boundaries delineating different levels of competence. In Study 3, participants learned about financial investors who all gained different levels of money (ordinal condition) or who crossed category boundaries of losing money, breaking even, and gaining money (categorical condition). In both studies, we hypothesized that participants would prefer to hire "relatively good" individuals only in the ordinal encoding conditions; in

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

the categorical condition, we predicted that this effect would be eliminated.

Finally, Study 4 tested our third prediction, asking whether a change in categorical gist would lead people to form different impressions, even if ordinal gist and verbatim details remain the same. Participants learned about two individuals who donated money to two charities. One individual donated a larger amount, but his donation was described in a manner that evoked the category "donated some" ("donated $45 to Charity A") or with redundant information that evoked the categories "donated some and none" ("donated $45 to Charity A and none to Charity B"). We hypothesized that participants would form worse impressions of this individual when categorical gist included "none," even when accounting for their verbatim memories of the amount donated. Altogether, these studies tested whether participants encode trait impressions in terms of categorical versus ordinal gist, and further, whether these two modes of social impression formation have distinct consequences for social decision-making.

Study 1a

In Study 1a, we asked whether ordinal impressions shape social decision-making under economic incentives, akin to judgments from memory and decisions rooted in reward-based learning (Palminteri & Lebreton, 2021; Sharif & Oppenheimer, 2016, 2021). Participants learned about two individuals who objectively were identically competent at winning points in a trivia game; however, one individual appeared among a group of more competent players and one individual appeared among a group of less competent players. We tested whether participants would later rescale their impressions to account for these distinct learning contexts or whether participants would maintain reference-dependent impressions, rating the "relatively competent" player as more competent and preferring that player in subsequent hiring decisions.

Method

Participants

We recruited 140 individuals in exchange for payment through Cloud Research, which draws on Amazon's Mechanical Turk platform (56 female, 82 male, two nonbinary; $M_{age} = 38.56$, $SD = 12.17$). This preregistered sample size was determined as follows. A pilot study yielded an effect size of $d = 0.35$ for the analysis of choice; to be conservative, we conducted a power analysis targeting 90% power while assuming a slightly smaller effect size ($d = 0.30$, $N = 119$) and increased this number to account for potential exclusions (15% anticipated, rounded to the total sample size of 140). Participants provided informed consent in accordance with approval from the University of Southern California (USC) Institutional Review Board.

Following preregistered exclusion criteria, data were excluded from analysis if a participant failed to respond on more than 20% of trials during either learning phase block (Kool et al., 2017) or if participants did not pass a hypothesis-independent learning criterion during the decision phase. Specifically, during decision phase trials that did not include the key targets, participants had to choose members of the "competent" group over members of the "incompetent" group in more than 50% of trials to be included in the analysis. This exclusion rule ensured participants had paid attention during the learning phase (i.e., they knew which targets had been competent or incompetent) in a manner independent of our hypothesis (i.e., this rule concerned only filler trials rather than trials involving the two key targets we used in our primary analyses). These criteria excluded data from 45 participants, leaving 95 participants for analysis. Although Cloud Research screens participants for quality, this exclusion rate was higher than expected based on our previous experience with this participant pool; however, subsequent studies had a lower exclusion rate in line with expectations.

Stimuli

Participants viewed avatars representing supposed previous participants in a "Worker" role. To avoid cues to social categories that could influence social judgments, Workers were represented by colorful animal avatars, similar to those used on prominent collaboration websites. The correspondence between the eight images and their feedback patterns (i.e., group type and competence level) was randomized across participants to control for any effects of a particular stimulus (e.g., pleasantness).

Procedure

Participants were told they would learn about previous participants ("Workers") who had answered trivia questions to win points worth money. Workers had supposedly been able to earn 100 points on each round, depending on how quickly and accurately they answered trivia questions.

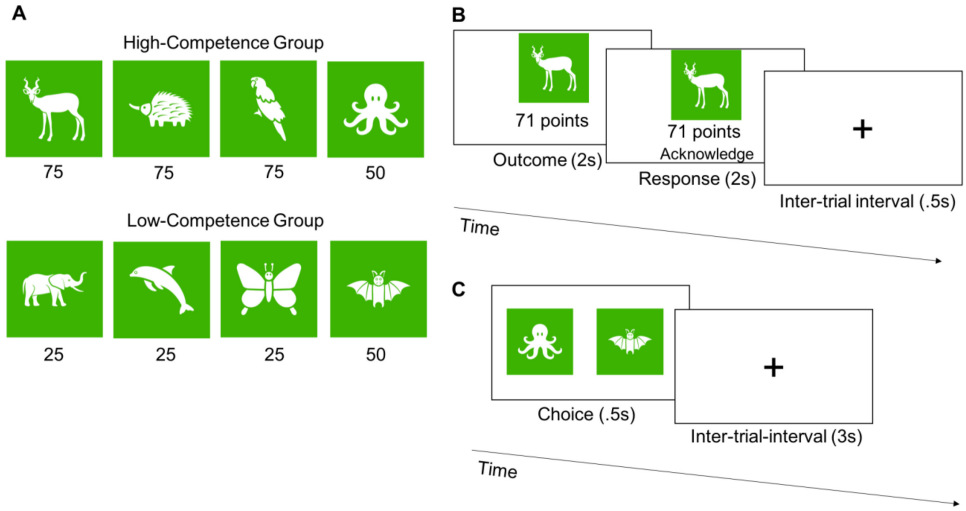
In the learning phase, participants saw the Workers' performance. Participants were told that they would learn about Workers in two groups that participated at different times, each of which had four workers. On each trial of the learning phase, participants saw a Worker's avatar appear onscreen along with the point score that Worker had won on that round out of 100 points; this information stayed onscreen for 3 s (Figure 1B). Afterward, participants had a 2 s window in which to press a button to acknowledge seeing the information. Participants encountered the groups in two separate blocks, each featuring 58 trials, with a brief rest break in between. Passive presentation of information was used to avoid reward-based learning, which could promote habit formation toward choosing a relatively rewarding target (Wood, 2017).

Unbeknownst to participants, one group had high competence (average points earned = 75) and one had low competence (average points earned = 25), except for one moderately competent key target in each group (average points earned = 50; Figure 1A). Point values on each trial were generated by using these average values plus Gaussian noise with $SD = 0.075$. Values followed a censored normal distribution, such that amount won had to be at least 2 points and had to be between 0 and 100. In this manner, all group members had point values drawn from a random normal distribution with means shifted. The instructions noted that targets in both groups saw the exact same set of trivia questions to forestall any inferences that questions faced by the "low-competence" group were harder. The order of groups (competent first or incompetent first) was randomized across participants.

To motivate learning, participants were told to pay careful attention because they would later make choices to "hire" different individuals, and their bonuses would depend on how well they learned. To allow participants to form expectations before seeing each

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Figure 1
Schematic of Task



Note. (A) Participants learned about two groups of four "Workers" who completed a trivia task for points worth money; Workers were represented by avatars. Values indicate the average number of points (out of 100) each Worker earned on each round. In each group, one Worker earned 50% of points, on average; however, these key targets were embedded among three lower-competence or higher-competence Workers. (B) On each round of the learning task, participants saw the outcome one Worker earned on one round and pressed a button to acknowledge the outcome. Each group appeared in a separate block of trials. (C) In a subsequent decision phase, participants made choices between pairings of Workers, including the two moderately competent Workers. See the online article for the color version of this figure.

group's key target, the order of targets was set such that the first 12 trials involved only filler group members rather than the key targets. Other than this rule, the order of pairings shown in each trial was randomized for each participant.

After learning about both groups, participants completed a decision phase in which they made additional choices without receiving feedback (Figure 1C). Participants were told that their task was to "hire" Workers to answer questions for them. Participants were further told that they would also receive whatever points had been earned by the Workers, and that points would be converted into a monetary bonus. The purpose of this phase was to reveal whether relative impressions formed *within* each group would drive preferences between identical targets when considered together. Participants, therefore, saw novel cross-group pairings of targets, viewing every possible combination of Workers from Group A matched with Workers from Group B two times (32 trials). Our primary question concerned trials in which participants chose between the two "moderately competent" key targets; the other trials served as filler trials (to avoid suspicion about the key trials) and as a baseline criterion for data exclusion (described above). Therefore, we included an additional 10 trials featuring the two "moderately competent" key targets paired together (for a total of 42 trials). Participants had 3 s to make each decision. No feedback was provided, to prevent further learning; instead, participants were informed that they would discover how they performed after they completed the task. This phase let us test whether participants were indifferent between the two key targets—who were equivalently competent—or preferred to hire the one who was relatively competent compared to their (incompetent) group.

Following the main task, participants rated each Worker's competence, allowing us to assess whether their explicit impressions also reflected relative performance. Ratings were made using a 7-point Likert-type scale (1 = *not at all*, 7 = *very much*). "Competent" was defined for participants as "intelligent, knowledgeable, and capable." Upon completing these measures, participants were paid a bonus based on the number of points they accrued.

Transparency and Openness

Information regarding our procedure for determining sample size, all data exclusions, all manipulations, and all measures included in this research are fully reported in this article. Materials, deidentified data, and analysis code are available at: https://osf.io/xwvue/?view_only=250a1baab4944912929afbdccc314880 (Hackel & Mende-Siedlecki, 2023). Preregis-tration documents are available at: https://aspredicted.org/4xh55.pdf (Study 1a), https://aspredicted.org/LJZ_2KL (Study 1b), https://aspredicted.org/7N8_GG6 (Study 2), https://aspredicted.org/FXG_663 (Study 3), and https://aspredicted.org/2PP_2X4 (Study 4).

Results

Decision Phase

Our primary question concerned participants' choices between the two key targets: were participants indifferent between these two identical, moderately competent Workers or did they prefer to hire the one who was relatively better than their (incompetent) group? To address this question, we fit participant choices from the decision phase to a

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

mixed-effects logistic regression predicting whether, on each trial featuring the two key targets, participants chose the one from the low-competence group (1 = yes, 0 = no). This model included a fixed intercept and a random intercept. In mixed-effects models, the "fixed" term models the average population effect (i.e., whether the average effect is significantly different from zero) whereas the "random" term models variability across subjects in the population (Brown, 2021). Accordingly, the fixed intercept in our model estimates whether participants chose the "relatively good" target more often than chance, on average—conceptually equivalent to asking whether the mean proportion of choices across subjects is significantly greater than .5. In contrast, the random intercept accounts for the fact that each participant completed many trials and that participants differed from one another (i.e., nonindependence of data points from the same participant). The model was fit using the lme4 package for R (Bates et al., 2014; R Core Team, 2016).

We hypothesized that participants would be more likely to choose the Worker from the low-competence group who appeared relatively competent; we, therefore, compared the intercept to zero, asking whether participants were more likely to choose this Worker than expected by chance. The intercept was significantly greater than zero, $b = 1.63$, $SE = 0.33$, $z = 4.96$, $p < .001$, indicating that participants indeed preferred to hire the relatively competent Worker over the relatively incompetent Worker. On average, participants chose the relatively competent Worker over the relatively incompetent Worker in 68% of trials (Figure 2A), 95% CI [0.61, 0.75], $d = .52$. Thus, when deciding whether to hire two equivalent candidates, participants preferred the one who had stood out positively from a low-competence group, even though both had behaved identically.

Explicit Impressions

We similarly examined whether participants formed explicit impressions that reflected relative, rather than absolute, competence—consistent with ordinal gist. Explicit ratings of competence for the two key targets were submitted to a paired-samples *t* test. Indeed, participants judged the "relatively competent" Worker to be more competent ($M = 4.66$, $SD = 1.24$) than the "relatively incompetent" Worker ($M = 3.80$, $SD = 1.24$), $t(94) = 4.68$, $p < .001$, 95% CI on mean difference [0.50, 1.23], $d = .48$ (Figure 2B). Participants thus formed more positive impressions of the "relatively competent" Worker, even though both Workers had performed identically during learning.

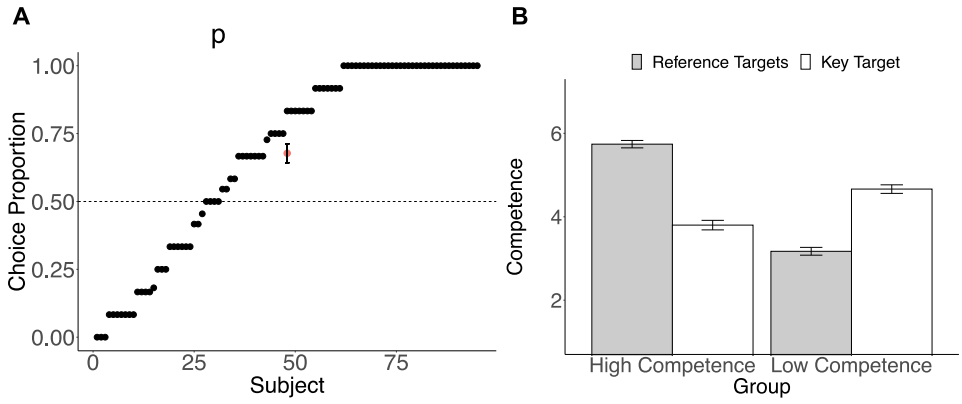
Discussion

In Study 1a, participants learned about two moderately competent individuals who appeared in groups with lower or higher average competence. In subsequent hiring decisions, participants preferred the individual who appeared relatively competent over the one who appeared relatively incompetent. Participants similarly formed more positive impressions of the former's competence, even though both individuals had identical performances. These findings suggest that participants encoded competence as ordinal gist and used ordinal impressions to make decisions, rather than relying on verbatim memories of performance (which would have been identical) or rescaling impressions to account for the different learning contexts.

Study 1b

In Study 1b, we asked whether ordinal impressions would generalize to impressions of generosity and prosocial decisions. Generosity reflects the dimension of "warmth," which forms an important dimension of person perception in addition to "competence" (e.g., Fiske et al., 2007) These impressions play a key role in prosocial decisions: people act more kindly to others who are generous (a phenomenon termed "indirect reciprocity"; Wedekind & Milinski, 2000). Yet, if impressions of generosity reflect ordinal gist, then indirect reciprocity may similarly reflect relative reference points: when faced with two equally generous individuals learned about in more or less generous groups, people might be kinder to the one encoded as relatively generous. Study 1b tested this hypothesis using a game in which participants learned about "givers" who shared money, after which they

Figure 2
Study 1a Results



Note. (A) Proportion of trials for which each participant chose the "relatively competent" moderate Worker over the "relatively incompetent" moderate Worker. The lone circle with error bars indicates mean across participants and *SEM*. The dotted line indicates chance. (B) Ratings of competence for each Worker after the task, including the moderately competent key targets; reference targets for each group are shown for comparison. Error bars represent *SEM*, adjusted for within-subjects comparison (Morey, 2008). See the online article for the color version of this figure.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

had the opportunity to allocate money to givers in return—a measure of indirect reciprocity.

Participants

We recruited 150 individuals through Amazon's Mechanical Turk (67 female, 83 male; $M_{age} = 39.19$, $SD = 10.38$) who participated in exchange for payment. This preregistered sample size was determined as follows. In Study 1a, the smallest effect size of interest was $d = 0.48$ (analysis of trait impressions). However, given that Study 1b involved participants making prosocial decisions rather than decisions yielding personal gain, we again recruited a larger sample in order to account for the possibility of a smaller effect size with this different design ($d = 0.30$, requiring $N = 119$) plus 25% additional subjects to account for potential subject exclusions, rounded up. (We did not expect the anomalously high exclusion rate from Study 1a to repeat, but we nonetheless increased the anticipated exclusions somewhat out of caution.) In Study 1a, our exclusion rule focused on decision phase trials featuring filler group members as a test of attention, given that participants were incentivized to choose the more competent group members. However, in Study 1b, choices in the decision phase depended on participants' subjective preferences for indirect reciprocity; a participant who did not value indirect reciprocity would not necessarily allocate money to others based on generosity, even if the participant had paid attention. We, therefore, presented participants with a separate attention check after the task, in which they saw pairs of background targets from each group and were asked to indicate which had shared a larger proportion of their points. We excluded participants who did not answer more than 50% of questions correctly. Using this exclusion rule, data from 27 participants were excluded, leaving 123 participants for analysis. Participants provided informed consent in accordance with approval from the USC Institutional Review Board.

Procedure

The procedures were adapted from those of Study 1a, with a few changes. First, the learning phase was framed as learning about generosity. Participants were told that they would learn about two groups of previous participants (Deciders) who had made decisions about sharing money with a future set of participants (Recipients). Supposedly, each Decider could share up to 100 points worth of money with the Recipients. Participants served in an "Observer" role, in which they witnessed how much Deciders shared with Recipients. Participants did not win money themselves but had an opportunity to witness the generosity of Deciders, as in prior studies of indirect reciprocity (e.g., Wedekind & Milinski, 2000).

On each trial of the learning phase, participants saw a Decider's avatar onscreen and the amount shared by that Decider out of 100 points. In one group, three Deciders shared 45% on average and one shared 30% on average. In a second group, three Deciders shared 15% on average and one shared 30% on average. The two key Deciders, therefore, shared an intermediate amount (30%), but they did so in a group context that made this amount appear relatively generous or relatively stingy. A small amount of Gaussian noise ($SD = 0.045$) was added to these values to render the task more interesting and less easy to learn; all group members thus shared proportions drawn from a random distribution with its mean shifted. Presentation order randomization and trial timings were the same as in prior studies. To motivate learning, participants were told to pay careful attention because they would later make choices that impact the individuals they are learning about.

In a subsequent choice phase, participants were told that they would now get to decide which of the "Deciders" they learned about would win additional points. They were told that this stage would be a surprise to the original Deciders, who did not know this would happen and who would receive an extra bonus based on participant decisions. On each round, participants saw two Deciders and chose to award 100 points to one of the two, allowing Deciders to accumulate more points over the course of many rounds. This phase therefore allowed the expression of indirect reciprocity based on prior impressions of generosity. The pairings of targets, timing, and all other elements of the design followed the choice phase of Study 1a.

After the task, participants rated each Decider's generosity, rather than rating their competence as in Study 1a. Next, participants completed an exploratory reciprocity phase in which they saw new behaviors and had the opportunity to share money with group members who acted more or less generous than their initial tendencies; this phase served as exploratory data for future work and is not discussed further here. Finally, participants completed an attention check to exclude nonattentive participants, described above.

Results

Indirect Reciprocity

Consistent with prior work (Wedekind & Milinski, 2000), participants engaged in indirect reciprocity overall: when choosing to award money to background members of each group, participants chose "generous" Deciders over "stingy" Deciders 87% of the time, which was significantly greater than chance levels of 50%, $t(122) = 23.35$, $p < .001$, $d = 2.11$. Participants thus learned about Deciders and rewarded Deciders according to their generosity.

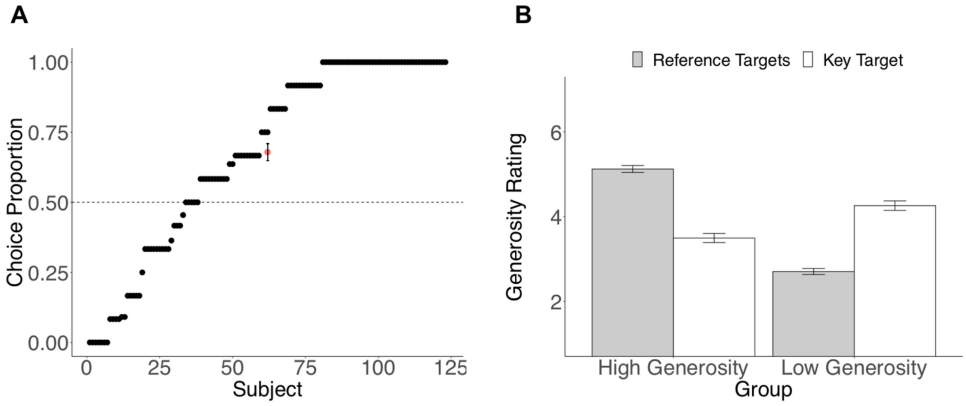
Beyond this overall preference toward indirect reciprocity, we tested whether participants preferred to reward the "relatively generous" key target over the "relatively stingy" one, even though both shared identical amounts on average. We again fit choices to a mixed-effects logistic regression, predicting whether, on each trial, participants chose the "relatively generous" target from the low-generosity group (1 = yes, 0 = no). This model again included a fixed intercept and random intercept; only trials featuring the two key targets were included in the model. The intercept was significantly greater than zero, $b = 1.62$, $SE = 0.29$, $z = 5.56$, $p < .001$. Participants chose the target from the low-generosity group in 68% of trials featuring these two targets (Figure 1), 95% CI [0.62, 0.74], $d = 0.53$ (Figure 3A). When deciding how to reward these two equivalent givers, participants thus preferred to allocate money to the one who had been relatively generous compared to a stingier group.

Explicit Impressions

We next directly tested whether participant impressions reflected ordinal encoding. Ratings of generosity were submitted to a paired-samples t test. As hypothesized, participants viewed the "relatively generous" Decider as more generous ($M = 4.26$, $SD = 1.52$) than

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Figure 3
Study 1b Results



Note. (A) Proportion of trials for which each participant chose the "relatively generous" moderate Decider over the "relatively stingy" moderate Decider. The lone circle with error bars indicates mean across participants and *SEM*. (B) Ratings of generosity for each Decider after the task. Error bars represent *SEM*, adjusted for within-subjects comparison (Morey, 2008). See the online article for the color version of this figure.

the "relatively stingy" Decider ($M = 3.50$, $SD = 1.31$), $t(122) = 4.01$, $p < .001$, 95% CI = [0.39, 1.14], $d = .36$ (Figure 3B).

Discussion

Study 1b replicated the findings of Study 1a in the domain of generosity, rather than competence. Participants learned about two individuals with equivalent generosity who appeared in more generous or less generous groups. In later choices between these individuals, participants preferred to allocate money to the "relatively generous" individual and rated this individual as more generous. Indirect reciprocity, therefore, reflected ordinal encoding of impressions, rather than each individual's objective generosity—consistent with the encoding of generosity in terms of ordinal gist.

Study 2

Studies 1a–1b demonstrated that people form ordinal impressions of competence and generosity that shape social decisions. In Study 2, we asked whether this tendency depends on the encoding style afforded to people by the task context. Fuzzy trace theory suggests that gist memories can reflect ordinal rankings or categories, and it predicts that people gravitate to the simplest representation of a task they can use. In Studies 1a–1b, participants lacked reference standards with which to interpret behavior; it is unclear what number of points in the trivia task reflects high or low competence or what number of points in the sharing game reflects high or low generosity. When perceivers lack reference standards with which to judge behavior, they may rely more on ordinal gist, forming impressions that compare one person to another. In contrast, when perceivers have reference standards they can use to categorize behavior, they may rely more on simpler categorical gist. In Study 2, we therefore manipulated the encoding style. Participants completed the same task as in Study 1a after random assignment to one of two encoding conditions. In the categorical condition, participants were given reference standards with which to interpret behaviors and were encouraged to categorize behaviors on each round; in the ordinal condition, participants were given no reference standards and were encouraged to compare targets to one another. We hypothesized that participants in the ordinal condition, but not the categorical condition, would prefer the "relatively competent" target.

Method

Participants

We recruited 200 individuals through Cloud Research (one nonbinary, 98 female, 101 male; $M_{age} = 41.08$, $SD = 11.53$) who participated in exchange for payment. This preregistered total was based on a power analysis aiming for 90% power to detect a moderate effect size ($d = 0.5$). This effect size was determined as follows: using the findings of Studies 1a–1b, we assumed a second group of subjects with a mean choice proportion of .50 (i.e., no effect of relative encoding) and the same standard deviations as subjects in Studies 1a–1b. We then computed the between-groups effect size that would be observed with these assumptions, which yielded an expected effect size of $d = 0.53$; we rounded this value down to be conservative. We again recruited 25% additional subjects to account for potential subject exclusions (rounded up to yield 200 total); we expected a lower number of exclusions that Study 1a, which had an anomalously high exclusion rate. One additional participant completed the study without requesting payment but did not make any responses during the choice phase and was therefore not included in analyses.

As in Study 1a, we excluded data from participants who did not correctly choose reference targets from the "competent" group over those from the "incompetent" group more than 50% of the time during the decision phase; we did not use the additional learning phase exclusion rule used in other studies because participants in this study did not have a time-locked response window for acknowledging information during the learning phase. The exclusion rule excluded data from 13 participants, leaving 190 participants for analysis. Four additional participants were missing explicit impression

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

ratings of one or both targets and were therefore excluded from analyses of explicit impressions.

Procedure

Participants learned about two groups of individuals who answered trivia questions to win points, using identical procedures to those of Study 1a; participants again learned about an "incompetent" group that earned an average of 25 points, except for one moderate target who earned 50 points on average, and a "competent" group with an average of 75 points, except for one moderate target who earned 50 points on average. Again, these point values varied on each trial with Gaussian noise, as in Study 1. However, participants were first randomly assigned to one of two conditions. In an "ordinal encoding" condition, participants were given no further instructions about the points earned. After every trial of the learning phase, participants were asked to rate that target relative to the rest of their group, using a scale ranging from −50 (*worse than others*) to 50 (*better than others*). In a "categorical encoding" condition, participants were given reference standards to use. Specifically, participants were told that Workers in this task tend to earn points in one of three ranges: 0–40 points ("poor"), 40–60 points ("moderate"), or 60–100 points ("excellent"). After each trial, participants categorized the target as poor, moderate, or excellent. As in Study 1a, participants completed a decision phase after learning in which we assessed decisions between the two key targets, followed by explicit ratings of target competence.

Results

Decision Phase

We examined choices between the two moderately competent key targets, asking whether participants in the ordinal condition—but not the categorical condition—preferred the key target who was relatively competent. We again fit participant choices from the decision phase to a mixed-effects logistic regression predicting whether, on each trial featuring the two key targets, participants chose the relatively competent Worker (1 = yes, 0 = no). We included the condition as a fixed effect predictor (1 = ordinal, −1 = categorical).

Encoding condition shaped decisions, as revealed by a significant effect of condition, $b = 0.76$, $SE = 0.26$, $z = 2.90$, $p = .004$ (Figure 4A). Simple effects analysis revealed that the ordinal encoding condition replicated the effects of Studies 1a–1b: the intercept was significantly greater than zero, $b = 1.62$, $SE = 0.39$, $z = 4.15$, $p < .001$, indicating that participants chose the relatively competent Worker over the relatively incompetent Worker more frequently than chance. In contrast, in the categorical encoding condition, the intercept was not significantly different from zero, $b = 0.09$, $SE = 0.36$, $z = 0.26$, $p = .79$, indicating no significant preference on average between these two workers. More concretely, participants in the ordinal encoding condition chose the Worker from the low-competence group in 65% of trials, on average, 95% CI [0.57, 0.73], $d = 0.41$, and participants in the categorical encoding condition chose the Worker from the low-competence group on 49% of trials, on average [0.41, 0.57], $d = -0.02$—a mean difference of 0.16, [0.05, 0.27], $d = 0.43$. Thus, a preference for a relatively positive target was eliminated when participants were given categorical reference standards to use during learning.

Explicit Impressions

Did encoding condition similarly shape explicit impressions? To test this question, competence ratings were submitted to a 2 (target type: relatively competent, relatively incompetent) × 2 (condition: categorical, ordinal) mixed ANOVA. A Target Type × Condition interaction indicated that the effect of target type depended on the condition, $F(1, 183) = 12.76$, $p < .001$, $\eta^2_p = .07$ (Figure 4B). In the ordinal encoding condition, participants viewed the "relatively competent" Worker as more competent ($M = 4.87$, $SD = 1.54$) than the "relatively incompetent" Worker ($M = 3.79$, $SD = 1.41$), even though both performed identically, $t(91) = 4.25$, $p < .001$, 95% CI on mean difference [0.57, 1.58], $d = 0.44$. In contrast, in the categorical encoding condition, participants showed no difference in impressions between the "relatively" competent worker ($M = 4.14$, $SD = 1.36$) and the "relatively incompetent" worker ($M = 4.23$, $SD = 1.33$), $t(92) = -0.42$, $p = .68$, [−0.49, 0.32], $d = -0.04$. Thus, biases in impressions were similarly eliminated by categorical reference standards.

Discussion

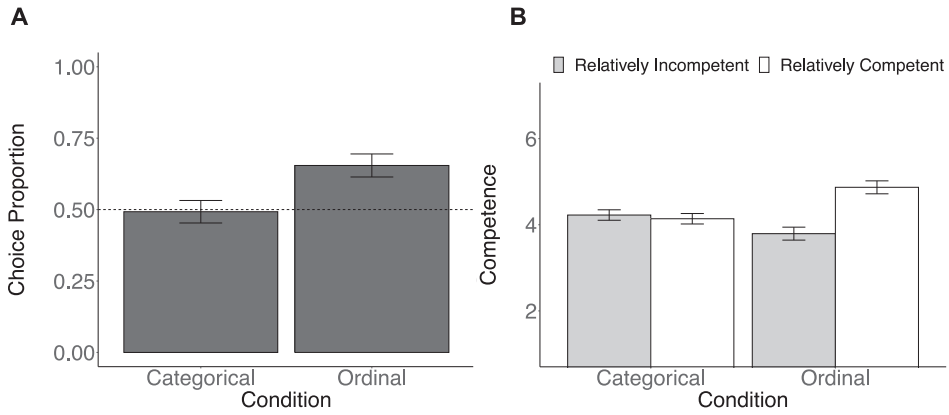
Study 2 demonstrated the impact of ordinal or categorical encoding of impressions on decision-making. Participants again learned about two individuals who were equally competent but who were embedded in more competent or less competent groups. In an "ordinal" condition, participants had no categorical reference standards with which to encode behavior; it was ambiguous whether "50 points" is inherently good or bad. Replicating Studies 1a–1b, participants in this condition preferred a "relatively competent" target over a "relatively incompetent" target and rated this target as more competent. In contrast, participants in a "categorical" condition were taught reference standards for categorizing behavior as "poor," "moderate," or "good." Participants in the categorical condition showed no significant preference or difference in ratings between the relatively competent and relatively incompetent targets. This finding suggests two modes of impression formation, corresponding to ordinal and categorical gist, and highlights their respective contributions to decision-making.

Study 3

In Study 3, we tested whether naturally occurring category boundaries would similarly eliminate preferences between these targets. Participants learned about the financial decisions of investors who had chosen stocks. In an "ordinal encoding" condition, incompetent investors had smaller gains, two moderate investors had intermediate gains, and competent investors had larger gains. In a "categorical encoding" condition, incompetent investors lost money, two moderate investors broke even on average, and competent investors gained money. Crucially, the objective differences between incompetent, moderate, and competent investors were held constant across conditions. However, in the ordinal condition, investors varied within a category (gains), promoting ordinal encoding, whereas in the categorical condition, investor performance crossed category lines (loss, break even, gain), allowing categorical encoding. We hypothesized that participants would gravitate to the simplest representation afforded by the task. Accordingly, participants in the ordinal condition, but not the categorical condition, would prefer to hire the

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Figure 4
Study 2 Results



Note. (A) Average proportion of trials for which participants chose the "relatively competent" Worker over the "relatively incompetent" Worker, separated by encoding condition. Error bars indicate *SEM*. Dotted line indicates chance. (B) Average posttask ratings of competence for the relatively competent and relatively incompetent Worker within each condition. Error bars represent *SEM*, adjusted for within-subjects comparison (Morey, 2008).

"relatively competent" investor over the "relatively incompetent" investor.

Method

Participants

A total of 240 participants were recruited through Cloud Research (102 female, one nonbinary, 137 male; $M_{age} = 39.89$, $SD = 11.82$) in exchange for payment. This preregistered total was based on a power analysis aiming for 90% power to detect an effect size of $d = 0.43$, as observed in the analysis of decisions in Study 2, plus 10 additional subjects to account for potential subject exclusions (using the exclusion rate of 7% observed in Study 2). We used the same exclusion rule described in Study 1a, which excluded data from 17 participants, leaving 223 participants for analysis.

Procedure

Procedures were adapted from those of Study 1a, except that the learning task was framed as one about financial competence. Participants were told they would learn about previous MTurk Workers who chose stocks ("Investors"). Supposedly, each Investor was given 40 stocks to research for a week prior to the study and then selected 15 stocks to "invest in" based on their research. For each stock, Investors were supposedly given 100 points to invest; after 6 months, they were paid a bonus based on how their stock had performed. If the stock went up, they received more than 100 points, and if the stock went down, they received less than 100 points. In the learning task, participants were told they would see the percentage that each stock had increased or decreased at the end of the 6 months. To motivate attention, participants were told that, in a second section, they would play a "Customer" role in which they choose which Investor to "hire" and win whatever the Investor earned on that round.

During the learning task, participants saw the percentage that an Investor's stock had increased or decreased on each trial; the Investor's avatar was shown onscreen along with the percentage, as in Study 1a. In the "ordinal" condition, investors in the incompetent group gained 10% on average, investors in the competent group gained 30% on average, and one moderate investor in each group earned 20% on average. In the categorical condition, investors in the incompetent group lost 10% on average, investors in the competent group gained 10% on average, and one moderate investor in each group gained/lost 0% on average. The difference between each set of values was thus identical across conditions (10 percentage points between each target type of incompetent, moderate, and competent), but the mean of the distributions varied such that the ordinal condition featured only gains and the categorical conditions featured losses, breaking even, and gains. Feedback during the learning phase was determined using these values above with Gaussian noise ($SD = 2.5$); in the ordinal condition, feedback was restricted to being positive (i.e., 1% or greater), to avoid crossing category boundaries. Feedback was displayed as a percentage value. Participants were not asked to make any categorization or comparison ratings during the learning phase; they simply observed the outcomes. The rest of the procedures were otherwise identical to those of Study 1a; after learning, participants completed a decision phase in which they chose between Investors and made explicit ratings of each Investor's competence.

Results

Decision Phase

We examined choices between the two moderately competent Investors, asking whether participants who saw only gains would prefer the "relatively competent" Investor whereas participants who saw outcomes in different categories would show no preference between the two key Investors. We again fit participant choices from the decision phase to a mixed-effects logistic regression predicting whether, on each trial featuring the two key Investors, participants chose the relatively competent one (1 = yes, 0 = no). We included

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

the condition as a fixed effect predictor (1 = ordinal, −1 = categorical).

Indeed, encoding condition-shaped decisions, as revealed by a significant effect of condition, $b = 1.21$, $SE = 0.24$, $z = 5.02$, $p < .001$ (Figure 5A). When all targets earned gains (ordinal condition), participants preferred the relatively competent Investor over the relatively incompetent Investor, $b = 2.45$, $SE = 0.37$, $z = 6.60$, $p < .001$. In contrast, when different Investors lost money, broke even, or gained money (categorical condition), participants showed no significant preference between these two Investors, $b = 0.03$, $SE = 0.31$, $z = 0.11$, $p = .91$. More specifically, participants in the "ordinal encoding" condition chose the relatively competent Investor on 72% of trials, on average, 95% CI [0.66, 0.79], $d = .64$, whereas participants in the "categorical encoding" condition chose this Investor on 48% of trials, on average, [0.41, 0.55], $d = -0.04$, reflecting a mean difference of 24%, [0.14, 0.34], $d = 0.66$ (Figure 1). Thus, a preference toward the relatively positive Investor was eliminated when Investors could be easily categorized.

Explicit Impressions

We similarly examined whether participants' explicit impressions reflected relative competence in the ordinal encoding condition but not in the categorical coding condition. Ratings of Investor competence for the two key targets were submitted to a 2 (target type: relatively competent, relatively incompetent) × 2 (condition: categorical, ordinal) mixed ANOVA. A main effect of condition, $F(1, 221) = 43.85$, $p < .001$, $\eta^2_p = .17$, indicated that participants in the ordinal encoding condition rated investors as being more competent on average, while a main effect of target type, $F(1, 221) = 24.37$, $p < .001$, $\eta^2_p = .10$, indicated that the "relatively competent" Investor received higher ratings. More importantly, a Target Type × Condition interaction indicated that the effect of target type depended on condition, $F(1, 221) = 19.12$, $p < .001$, $\eta^2_p = .08$ (Figure 5B). In the ordinal encoding condition, we observed the same pattern of impressions found in Study 1a: participants rated the "relatively competent" Investor as more competent ($M = 5.01$, $SD = 1.42$) than the "relatively incompetent" Investor ($M = 3.86$, $SD = 1.42$), even though both performed identically, $t(107) = -5.81$, $p < .001$, 95% CI on mean difference = [0.76, 1.54], $d = 0.56$. In contrast, in the categorical encoding condition, participants showed no difference in impressions between the "relatively competent" Investor ($M = 3.65$, $SD = 1.11$) and the "relatively incompetent" Investor ($M = 3.58$, $SD = 1.24$), $t(114) = 0.46$, $p = .65$, [−0.23, 0.37], $d = 0.04$. Thus, reference-dependence in impressions was similarly eliminated by categorical encoding.

Discussion

Study 3 replicated the findings of Study 2 in a new domain (financial decisions) and using a more naturalistic manipulation: participants learned about investors who either gained different amounts of money (ordinal condition) or who lost, broke even, or gained money (categorical condition). According to fuzzy trace theory, people use the simplest representation of a task available to them (Corbin et al., 2015). For participants in the categorical condition, the simplest way to track investors would be to remember whether they had lost, gained, or broken even on average. For participants in the ordinal condition, this simple representation was unavailable; they therefore had to use a more detailed representation of ordinal ranking. Consistent with this view, participants in the ordinal condition preferred to hire the "relatively competent" investor over the "relatively incompetent" investor and rated the former as more competent than the latter. In contrast, participants in the categorical condition showed no such preference.

Study 4

In Studies 2–3, adding reference categories to a task changed impressions, as participants gravitated to a simpler categorical representation rather than relying on ordinal encoding alone. In Study 4, we asked whether *changing* the categorical gist participants used to encode behavior would change their impressions and decisions. We also directly measured verbatim memory, allowing us to test a role for gist above and beyond memory for exact details. To change the categorical gist, we again drew on fuzzy trace theory, which predicts that redundant information can change the categorical gist by evoking a new category. For instance, in the dread disease problem (Kahneman & Tversky, 1984), specifying "a 2/3 probability that everyone dies *and a ⅓ probability no one dies*" adds no meaningful information but evokes the category "none die"; removing this information therefore changes people's preferences (Reyna et al., 2014). Drawing on this logic, participants in Study 4 learned about two individuals who had donated to fundraisers; one individual donated a larger total amount to one charity while the other individual donated a smaller amount spread across two charities. We manipulated whether the individual who had donated more was described as "donating $45 to Charity A" or as "donating $45 to Charity A and none to Charity B." Although both conditions described the same behavior, the latter condition included redundant information that evoked the category "none." We hypothesized that participants exposed to the category "none" would form more negative impressions of this individual and would be less willing to help him, even when accounting for their memory for verbatim details.

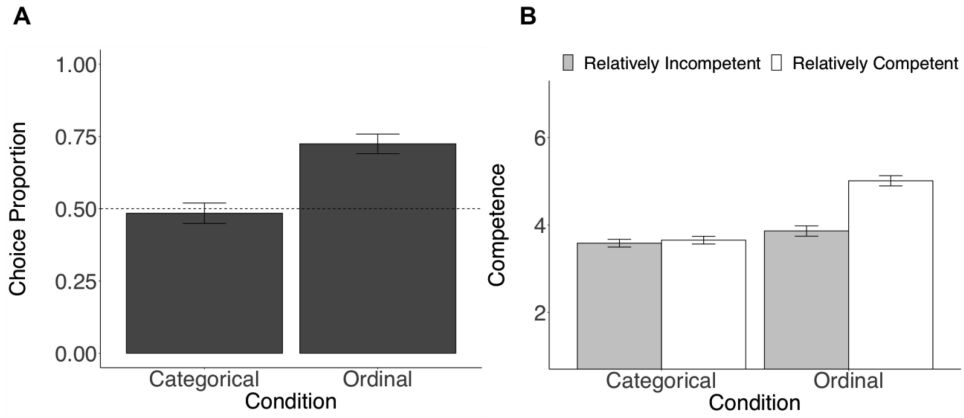
Method

Participants

A total of 666 participants were recruited through Cloud Research (334 female, 237 male, three nonbinary, and two identified as transgender men; $M_{age} = 42.16$, $SD = 12.91$) in exchange for payment. This sample size was based on a simulation-based power analysis aiming for 90% power for analyses of trait impressions based on a pilot study, which yielded a sample size of 600 participants, plus additional participants to account for potential subject exclusions (10% anticipated based on the pilot study). Though we preregistered a total of 660 participants to be recruited, an additional six participants completed the study without payment. Given the different design of Study 4 relative to Studies 1–3, we used a new (preregistered) exclusion rule. At the end of the experiment, subjects were asked what kind of scenario they had learned about and were presented with seven options (including "A night for parents at their children's school"). Participants were excluded from the analysis if they did not answer correctly. This rule excluded data from 70 participants, leaving 596 participants for analysis. An additional 10 participants did not respond to at least one of the verbatim memory questions about the exact amount an individual donated and were

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Figure 5
Study 3 Results



Note. (A) Average proportion of trials for which participants chose the "relatively competent" Investor over the "relatively incompetent" Investor, separated by encoding condition. Error bars indicate *SEM*. (B) Average posttask ratings of competence for the relatively competent and relatively incompetent Investors, separated by condition. Error bars represent *SEM*, adjusted for within-subjects comparison (Morey, 2008).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

excluded from analyses of these responses (but included in other analyses).

Procedure

Participants were told that they would learn about two individuals, John and Alex, each of whom has children attending the same school. Participants were told that the school held an event for families that included fundraisers for two charities—Charity A and Charity B—and were told they would learn some information about each individual's night. Specifically, participants were told they would learn which side of town each person came from, the total amount each person donated across the two charities and which charities they gave to, and what grade each person's child was in. The information about the side of town and the child's grade served as filler information to distract from the main purpose of the experiment. Importantly, however, the instruction about the charities made clear that participants would see the full amount donated by each individual (i.e., any and all donations would be displayed).

Participants then saw each piece of information about John, with one piece of information displayed onscreen at a time, followed by each piece of information about Alex; each individual was represented by an avatar similar to those used in prior experiments, and participants could press a button to proceed after they read each piece of information. However, participants were randomly assigned to one of two conditions. In a "None-Absent" condition, participants learned that John drove from the northwest side of town, donated $45 to Charity A, and had a child in second grade; next, they learned that Alex drove from the southwest side of town, donated $20 to Charity A and $20 to Charity B, and had a child in third grade. In a "None-Present" condition, participants explicitly read that John donated $45 to Charity A and none to Charity B; all other text was the same. Although this additional clause provides no new information about John's donation, it evokes the category "none," changing the categorical gist from "John donated some" to "John donated some and none." We hypothesized that this change would worsen impressions of John, even though he had donated more money than Alex.

After learning this information, participants rated each individual's generosity on a sliding scale ranging from 0 (*not at all generous*) to 100 (*very generous*). Next, participants made prosocial decisions involving John and Alex. They were told the school held a raffle for different prizes that night, and they were asked to imagine they had nine raffle tickets available to give away to help someone win a prize. They were asked how many of these tickets they would give to John and how many they would give to Alex. We assessed the number of tickets they assigned to John versus Alex as a measure of prosocial preferences.

Finally, in a planned exploratory phase (which was described as such in our preregistration), we assessed participants' verbatim memory of the donations. First, participants were asked how many charities each individual donated to; for each individual, participants could answer "One," "Two," "I don't know," or "This information wasn't provided." This measure let us determine whether participants fully understood the information they had seen, regardless of condition. Specifically, we verified that participants in the "None-Absent" condition understood that John had *not* donated to Charity B, even though this information was conveyed by omission instead of explicitly stated. Second, participants were asked to indicate the exact amount that each individual had donated to each of the two charities, using sliding scales ranging from $0 to $100. This measure allowed us to test whether gist condition would impact impressions even when statistically adjusting for verbatim memory.

Results

Prosocial Decisions

We first examined how participants divided raffle tickets between John and Alex. Given that John donated slightly more than Alex in total ($45 vs. $40), theories of indirect reciprocity would predict that

participants slightly prefer helping John. However, we hypothesized that this tendency would vary across conditions, such that participants who saw the category "none" would have a weaker preference for John. We anticipated in our preregistration that most subjects would give five tickets to one individual and four to the other, which was indeed the case; 93% of participants had a 5–4 split. We, therefore, classified participant choices as a binary outcome indicating which individual they chose to give more tickets. We compared the proportion of participants favoring John across conditions using a z test of proportions.

The proportion of participants who favored John (the greater donor) varied across conditions, $z = 2.30$, $p = .02$ (Figure 6A). When participants merely saw how much money John donated to Charity A ("None-Absent" condition), 58% of participants gave more raffle tickets to John, 95% CI [0.52, 0.63]—a proportion significantly greater than 50%, $z = 2.66$, $p = .008$. This small preference for John is consistent with indirect reciprocity, given that John donated slightly more money. In contrast, when participants explicitly saw that John donated none to Charity B ("None-Present" condition), they had no preference for John over Alex, with only 48% of participants favoring John, $z = .58$, $p = .56$, [0.43, 0.54]. Thus, participants preferred to help John when the categorical gist included only "donated some" but not when it included "some and none," even though he donated the same amount in both cases.

Explicit Impressions

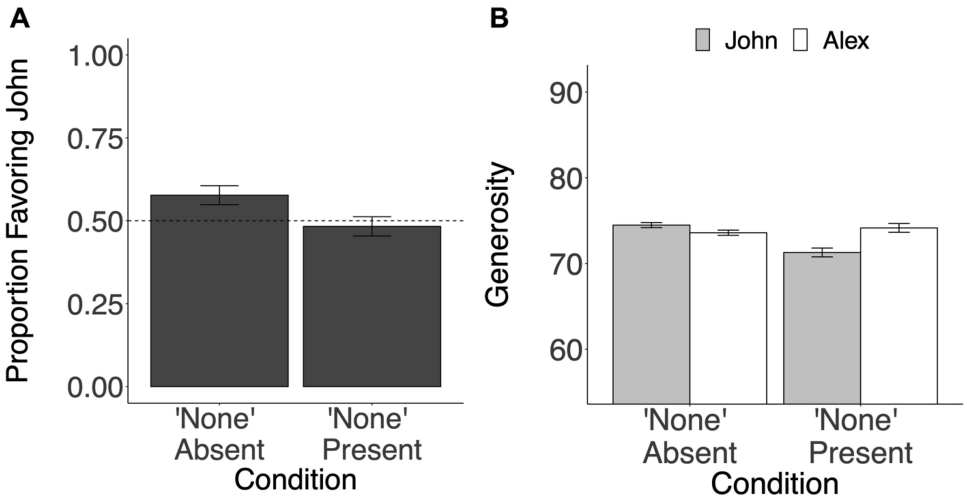
We similarly examined whether participants had worse explicit impressions of John when categorical gist included the category "none." Ratings of generosity were submitted to a 2 (target: John, Alex) × 2 (condition: none-absent, none-present) mixed ANOVA, with target as a repeated measure and condition as a between-subjects factor. A Target Type × Condition interaction indicated that relative ratings for the different targets depended on condition, $F(1, 594) = 14.05$, $p < .001$, $\eta^2_p = .02$ (Figure 6B). In the "None-Absent" condition, participants rated John ($M = 74.48$, $SD = 17.33$) as slightly more generous than Alex ($M = 73.59$, $SD = 19.07$), although this difference was not statistically significant, $t(297) = 1.76$, $p = .08$, 95% CI on mean difference = [−0.10, 1.89], $d = 0.10$. In contrast, in the "None-Present" condition, participants rated John ($M = 71.29$, $SD = 18.24$) as *less* generous than Alex ($M = 74.15$, $SD = 17.18$), $t(297) = -3.31$, $p = .001$, [−4.56, −1.16], $d = -0.19$. Thus, changing the categorical gist also changed social impressions: John appeared slightly more generous than Alex with a categorical gist of "some" but appeared less generous than Alex when categorical gist included "none."

Verbatim Memory

In exploratory analyses, we examined verbatim memory for donation details, for two reasons. First, this allowed us to consider only participants who interpreted the instructions correctly regardless of condition, understanding that John only gave to Charity A. Indeed, when asked how many charities each individual had donated to, 90% of participants in each condition correctly indicated that John had donated to one charity. When restricting our analyses to participants who correctly answered that John donated to one charity and Alex donated to two charities ($N = 506$), the findings reported above remained significant for indirect reciprocity, $z = 2.33$, $p = .02$, and for trait ratings, $F(1, 504) = 14.08$, $p < .001$, $\eta^2_p = .03$.

Second, we tested whether the effects of gist memory persisted when accounting for verbatim memory. According to fuzzy trace theory, people form verbatim and gist memories in parallel, and gist memories influence decisions above and beyond verbatim memory. Indeed, participants across conditions remembered similar amounts donated to each charity by each individual (Table S1 in

Figure 6
Study 4 Results



Note. (A) Proportion of participants who gave more raffle tickets to John (the individual who donated more overall), separated by encoding condition. Error bars indicate *SEM*. The dotted line indicates chance. (B) Average post-task ratings of generosity for each of the two donors, separated by condition. Error bars represent *SEM*, adjusted for within-subjects comparison (Morey, 2008). To facilitate intuitive estimation of effect size, the y-axis range includes 1.5 *SD*s of the ratings (Witt, 2019).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

the online supplemental materials) and the condition had no significant interactions with the target or charity when examining verbatim memory ($ps > .38$; Table S2 in the online supplemental materials). To test the roles of verbatim and gist memory more directly, we computed a difference score of verbatim memory, indicating the total amount participants remembered John donating across both charities minus the total amount they remembered Alex donating. We analogously computed a different score of impressions, indicating generosity ratings for John minus generosity ratings for Alex. We regressed the impression score on the verbatim memory score (standardized to z scores) and gist condition ("None-Absent" = 1, "None-Present" = −1). We observed effects of both verbatim and gist memory: participants rated John as relatively more generous than Alex to the extent that they recalled him giving a larger total amount, $b = 2.38$, $SE = 0.50$, $t(583) = 4.76$, $p < .001$, but they also rated him as relatively more generous if they had not encountered the category "none," $b = 1.87$, $SE = 0.50$, $t(583) = 3.74$, $p < .001$. A similar pattern emerged for indirect reciprocity: in a logistic regression, even when adjusting for verbatim memory ($b = 0.14$, $SE = 0.09$, $z = 1.65$, $p = .10$), participants were more likely to give more tickets to John in the "None-Absent" condition than the "None-Present" condition, $b = 0.19$, $SE = 0.08$, $z = 2.27$, $p = .02$. Altogether, these results identify an effect of gist above and beyond verbatim memory: although participants encoded similar details in verbatim memory regardless of condition, their impressions and decisions differed across conditions nonetheless.

Discussion

Study 4 found that changing a perceiver's categorical gist changed their impressions and decisions, even when verbatim details of a behavior remained constant. When participants read that John donated $45 to Charity A and Alex donated $20 to each of the two charities, they formed more positive impressions of John and preferred to help him over Alex. This tendency is consistent with indirect reciprocity, given that John had donated a larger total amount. However, when participants read that John donated $45 to Charity A *and none to Charity B*—changing categorical gist from "gave some" to "gave some and none"—they formed worse impressions of John than Alex and had no preference to help John. These findings held true even though the vast majority of participants in both conditions understood that John gave to only one charity and even when statistically accounting for participants' verbatim memories of the amounts donated. Altogether, these findings support an account of social impressions and choices rooted in fuzzy trace theory: people simultaneously encode verbatim details and gist, and moreover, changing categorical gist can change their choices for better or for worse.

General Discussion

Beyond remembering exactly what others have said and done, we remember the gist of their behavior: we remember that a colleague who comforts us was kind, a student with an excellent paper was competent, and a comedian with a sharp act was entertaining. These impressions persist in our memory even after we forget the details of their comforting words, brilliant arguments, or wry observations about airplane food (Falvello et al., 2015; Todorov & Uleman, 2002). The present work suggests that these spontaneous impressions reflect "fuzzy traces" of social behavior that can be formed in two modes—ordinal and categorical gist—with distinct consequences for social decisions. Across four experiments, participants learned about two individuals who displayed equivalent generosity or competence but who were embedded in groups that made them look relatively good or relatively bad. When participants had no reference standards with which to categorize behavior, they preferred to hire or share with the "relatively good" individual. In contrast, when participants could categorize behaviors based on preexisting reference standards, this preference was eliminated; ironically, a coarser, categorical representation thus led to more accurate judgments. In a fifth experiment, participants learned about two individuals who donated money to charity. All participants encoded details of the donations similarly, but when one individual evoked the category "donated none," participants formed worse impressions of him and had a weaker preference to help him. Altogether, gist shaped impressions and choices both when participants faced economic incentives (Studies 1a, 2, and 3) and prosocial incentives (Studies 1b and 4).

These findings link spontaneous trait impressions to fuzzy trace theory, dissociating modes of impression formation in memory and their consequences for behavior. According to fuzzy trace theory, people simultaneously form verbatim memories for details of an event and gist memories of essential meaning, and gist representations persist in memory past verbatim ones (Brainerd & Reyna, 2001; Reyna & Brainerd, 1995). Similarly, people encode both details of others' behavior and impressions of others' traits, and trait impressions persist past verbatim details in memory (Falvello et al., 2015; Todorov & Uleman, 2002, 2003; Winter & Uleman, 1984). Fuzzy trace theory further suggests that gist memories are part of a spectrum including ordinal and categorical gist, which vary in the level of detail they include from ordinal rankings to simple categories (Corbin et al., 2015; Mills et al., 2008). The present work suggests that social impressions can similarly be formed in an ordinal scheme ("more competent," "less competent") or categorical scheme ("poor competence," "moderate competence," "excellent competence"). Traditional measures of social impressions, as well as the hiring and helping behaviors they give rise to, are thus well-predicted by the theoretical variables of fuzzy trace theory.

This insight complements past work asking how social cognition shapes gist representations. According to fuzzy trace theory, people make decisions by applying values retrieved from memory ("saving some lives is better than none") to gist representations of new problems ("some lives will be saved"; Reyna, 2012a, 2012b), and social norms and social knowledge shape the values people hold and the gist representations they form (Reyna, 2012a; Rivers et al., 2008). Conversely, the present work asks how people form gist representations of social behavior itself, thus adding to recent work bridging fuzzy trace theory with impression formation. In particular, Sharif and Oppenheimer (2021) asked participants to read the resumes of two groups of job candidates, with two moderate resumes embedded in better or worse groups. Participants later preferred the "relatively good" candidate, but this preference was eliminated when participants were told that the two groups reflected students from different universities (i.e., two different categories). Complementing these findings, the present findings indicate that the content of the trait impression itself can reflect ordinal gist ("more competent," "less competent") or categorical gist ("competent," "incompetent").

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

In turn, these findings offer an expanded understanding of how different memory representations serve social decision-making. Social behavior depends on multiple memory systems, including semantic association, Pavlovian conditioning, and instrumental learning from rewards (Amodio, 2019). Recent work has begun to examine how episodic memory gives rise to social decisions, probing the role of exact representations of social encounters (FeldmanHall et al., 2021; Murty et al., 2016). For instance, when people remember the exact amount of money another person shared with them, they are more likely to choose that person again for an economic game if the amount shared was high (Murty et al., 2016). However, memories of events can be encoded not only as verbatim representations but also as abstract gists, and three lines of reasoning suggest gists may play a particularly strong role in social decisions. First, gists are longer-lasting in memory than verbatim details (Corbin et al., 2015), and this principle holds true for social impressions (Falvello et al., 2015); patients with amnesia have been found to recall good or bad impressions even when they cannot explicitly remember another person's identity (Johnson et al., 1985; Todorov & Olson, 2008; Tranel & Damasio, 1993). Second, people rely more on gist with increasing expertise, given that gists allow meaning-making (Corbin et al., 2015; Reyna et al., 2014). Humans are social experts with a wealth of highly structured social knowledge, able to easily recognize abstract meaning in social behaviors (Fiske, 1993; Hackel et al., 2022; Kalkstein et al., 2020; Read, 1987; Todorov & Uleman, 2002; Winter & Uleman, 1984). Third, when learning specifically about other humans (as opposed to nonsocial targets), people make social decisions primarily based on the abstract traits others display, rather than the concretely rewarding outcomes others provide (Hackel et al., 2020).

Although Studies 1–3 did not directly examine verbatim memory for details of specific instances (e.g., asking participants exactly how much money a target shared or earned), these studies highlight a role for gist that cannot be explained by verbatim memory; participants using verbatim memory would never have preferred the "relatively good" target over the "relatively bad" target, since the objective details of encounters with each target were the same. In Study 4, we found that participants reported similar verbatim memory across conditions but nonetheless reported different impressions, and the effects of categorical gist held when statistically adjusting for verbatim memory. These findings fit with the view that people form verbatim and gist memories in parallel, and that gists can influence decisions above and beyond verbatim memories (Reyna & Brainerd, 1995). Nonetheless, further examining the relative contributions of verbatim and gist memory to social decisions, the timescales upon which each type of representation influences decision-making, and reliance on gist across social versus nonsocial scenarios present interesting directions for future research. Altogether, the present findings expand models of memory representation in social decision-making, highlighting a distinction between verbatim episodic memory and gist, with two types of gist bearing distinct consequences for social decisions.

Implications for Social Behavior

The present findings may help predict when people do or do not show reference-dependent preferences in social decisions. Fuzzy trace theory posits that people gravitate to the simplest representation that will be useful for a given task or context (Corbin et al., 2015). When clear reference standards exist for interpreting behavior and a perceiver has considerable expertise with these standards, the simplest representation may be categorical, leading to less fine-grained choices that adhere to category boundaries. In contrast, when no clear reference standards exist, the simplest representation may be ordinal, leading to choices rooted in relative performance. Finally, when people are confronted with a truly novel behavior for which they cannot extract abstract meaning at all, they may rely more on verbatim memory, leading to choices rooted in exact behavior. These forms of gist are not mutually exclusive, as people can encode events in multiple ways at once (Brainerd & Reyna, 2001; Corbin et al., 2015). It is also possible that cultural background might influence encoding, given differences across cultures in holistic versus analytic attention, with holistic attention focusing on relationships between people and analytic attention focusing on abstract features or categories linked to one person (Nisbett et al., 2001; Nisbett & Miyamoto, 2005).

Although ordinal encoding led to seemingly biased preferences in the present work, we do not intend to suggest that one mode of encoding is inherently better or that categorical encoding always "debiases" social judgments; these modes may allow functional behavior in different settings. For instance, although categorical gist did not lead to reference-dependent preferences here, ordinal gist maintains more fine-grained distinctions than categorical gist, which may be important for many decisions—and in some situations, categorical encoding might also produce reference-dependent decisions (see Constraints on Generality). Moreover, in Study 4, changing participants' categorical gists led them to make decisions that were *less* consistent with indirect reciprocity, such that they no longer preferred to help an individual who had donated a larger amount. Altogether, the key feature of categorical gist is that it discretizes a continuous space, not that it renders decisions more objective. Nonetheless, the present work suggests that different representations of social behavior lead to different patterns of decision-making, for better or for worse.

In doing so, these findings also offer new insight into the mental representations underlying indirect reciprocity. People are more willing to pay a cost to help others who have themselves been generous (Wedekind & Milinski, 2000). This tendency encourages kindness to grow, given that people become more generous when they believe others are watching and can reciprocate (Barclay & Willer, 2007; Nowak & Sigmund, 1998; Wedekind & Braithwaite, 2002; Yoeli et al., 2013). Past work has typically examined indirect reciprocity by manipulating or measuring the actual generous behavior of targets (e.g., Hackel & Zaki, 2018; Sommerfeld et al., 2007; Wedekind & Milinski, 2000). However, indirect reciprocity depends not only on a giver's behavior but also on a perceiver's mental representation of that behavior. The present work demonstrates that perceivers use gist representations in indirect reciprocity rather than using verbatim recollection alone. These gist representations can lead people to prefer helping an individual encoded as "relatively generous" over one encoded as "relatively selfish," even when both individuals displayed objectively identical generosity, or can lead people to prefer helping an individual who donated a smaller amount to two charities (a gist of "some") over an individual who donated a larger amount to one charity (a gist of "some and none").

In the present work, ad hoc groups served as a reference point. Indeed, in daily life, people may form and update reference points

based on recent experiences (Dotsch et al., 2017), with different reference points forming in different contexts to reflect distinct experiences (Hunter & Gershman, 2018). At the same time, given our use of avatars to represent the social targets in these studies, our participants were ultimately learning about "raceless, genderless" individuals (Hester & Gray, 2020). While this approach is often used in psychological research to determine the basic contours of a phenomenon, identity factors likely serve as important moderators of impression formation processes studied here. Indeed, in other situations, existing social groups may serve as reference points, such that individuals are encoded as deviations from expectations or stereotypes about their group. For example, neuroimaging work on the development of novel group-based stereotypes demonstrates that group members whose behavior diverges from their group's tendencies elicit prediction error signals in the anterior temporal cortex, dorsomedial prefrontal cortex (PFC), and lateral PFC (Spiers et al., 2017)—brain areas involved in social impression formation and updating (Mende-Siedlecki, 2018; Mende-Siedlecki, Cai, & Todorov, 2013). However, in the context of real-world social groups, the stereotype-inconsistent behavior of individual group members can be "subtyped," separating the individual from their group as an "exception to the rule" and consequently maintaining the group stereotype (Hewstone, 1994; Richards & Hewstone, 2001). Separately, group stereotypes could also be strong enough to guide the encoding of behavior, in turn, confirming those stereotypes (Darley & Gross, 1983). Moreover, groups can also lead to assimilation effects, in which an individual is evaluated similarly to the group, depending on the connection between an individual and other group members (Bless & Schwarz, 2010). Ultimately, future research can examine how categorical representations of traits interface with contrast effects in preexisting social groups.

Finally, while the present work considered the consequences of reference-dependent social impressions in relatively circumscribed paradigms, subsequent research could examine this phenomenon in real-world decision contexts. Hiring, admissions, and financial aid decisions all represent likely candidates for domains in which reference-dependence can exert an influence. Indeed, related work has recently demonstrated how social decoys shift preferences in hiring decisions, how this effect varies by context (e.g., whether trait warmth or competence is prioritized), and how this effect can be harnessed to influence bias in hiring decisions (Chang & Cikara, 2018). In addition, electoral decision-making marks another potentially fruitful avenue for further inquiry. Indeed, some work has already considered how this phenomenon might impact voter choice in the context of political primaries (Bendle, 2014), where candidates across parties might vary in terms of ideological extremity, and thus, electability.

Implications for Learning, Judgment, and Decision-Making

A growing body of research has demonstrated context-dependence in nonsocial reward learning, similar to the reference-dependence in social choices observed here (Palminteri & Lebreton, 2021). For instance, after learning about a "relatively bad" slot machine in a gain context and a "relatively good" slot machine in a loss context, people ironically prefer the "relatively good" slot machine in choices between the two—even though it is objectively worse (Palminteri et al., 2015). Computational models suggest that people track the average rewards available in a context and encode rewards relative to that baseline.

Notably, decision-making in reward-learning tasks can reflect episodic memory for the details of an outcome in addition to a cached value reflecting the long-term average reward acquired from an action (Biderman & Shohamy, 2021; Bornstein et al., 2017; Murty et al., 2016). This distinction resembles the distinction between verbatim and gist representations in fuzzy trace theory, raising the question of whether reward value may also be encoded categorically. In the present work, preexisting categories shaped how participants encoded outcomes used to make economic choices. On the other hand, the present work involved passive learning reliant on declarative memory, whereas reward learning involves active instrumental learning from choice and feedback. These two types of learning reflect distinct neural systems with distinct influences on behavior (Amodio, 2019; Foerde et al., 2006; Poldrack et al., 2001). It is possible that instrumental learning, unlike declarative memory, encodes rewards in a context-dependent scheme regardless of category knowledge. Alternatively, it is possible that categorization influences nonsocial reward learning, much as other forms of semantic knowledge or categorization can influence affective learning and experience (Doll et al., 2009; Dunsmoor & Murphy, 2014; Satpute et al., 2016). Future work can test this potential interface between categorization and reward learning, examining the boundaries of the present findings and the scope of context-dependence in economic decision-making.

Finally, in the interest of clarifying the conceptual scope of the current work, we highlight some important distinctions with other phenomena that may also seem related to the present findings. First, people often engage in "narrow bracketing," treating a small sample of a choice set like the full set; for instance, they might advance 20% of job candidates on a given day if they aim to advance 20% on the whole (Simonsohn & Gino, 2013). The present work similarly finds that people group individuals into subsets when forming judgments and choose individuals based on how they fit into that subset. However, research on narrow bracketing concerns decisions within one subset, whereas the present findings concern how people form mental representations in different modes (i.e., on an ordinal or discrete scale) that shape their decisions between individuals from different subsets.

Second, the present findings relate to decoy effects—another form of context-dependence in which adding a third option changes people's preferences between two initial options (Chang & Cikara, 2018; Huber et al., 1982; Simonson, 1989). Unlike decoy effects, however, the present findings concern the granularity with which people encode *one* dimension of social impressions (e.g., warmth), rather than compromises people make along two dimensions (e.g., warmth and competence).

Third, when employers use a standardized process for hiring—for instance, holding constant the features they evaluate in candidates or the weight they give each feature—their decisions become more reliable, consistent, and predictive (Campion et al., 1988, 1994; Dawes, 1979). Study 2 of the present work highlights how categorical gist can sometimes standardize judgments within a single feature: when participants used explicitly defined cutoffs for "poor," "moderate," and "excellent" performance, they made decisions that more objectively reflected each worker's performance. However, categorical gist is not synonymous with standardization or "debiasing": in Studies 3–4, participants used intuitive categories without any

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

standardized instructions, and in Study 4, categorical gist led participants to make choices that were *less* reflective of each giver's donations. More generally, fuzzy trace theory posits that categorical gist is context-dependent, reflecting the way an individual makes meaning of events around them (Kühberger & Tanner, 2010; Reyna et al., 2021). Categorical gist thus can overlap with but does not inherently involve, standardization.

Constraints on Generality

Two features of the current studies are worth noting as potential boundary conditions. First, participants in the present studies saw behaviors reflected in numbers. Indeed, people often form social impressions from numbers: How much did a colleague tip at the restaurant? How late did a worker arrive at the office? What grade point average did a graduate applicant have? What returns did a financial advisor achieve? How many fully loaded nachos with meat and cheese did a date eat from the shared dish, leaving behind bare chips? Past work in social psychology has often left out these numerical details, instead focusing on abstract secondhand descriptions of behavior (e.g., "John gave the waiter a low tip"). Yet, in firsthand experience, people see specifics rather than summaries, and they must draw their own qualitative inferences from quantitative data. The present work informs how people reduce rich details ("he tipped 5%") into social meaning (e.g., "he tipped less than my last date" or simply "he was a cheapskate"). While an open question concerns whether these findings would emerge for other kinds of information, recent work is encouraging on this point. People do form relative memories for nonnumeric information (e.g., evaluating CVs or singing auditions; Sharif & Oppenheimer, 2016, 2021), and they might similarly form categorical impressions from this information (e.g., "college graduate" vs. "noncollege-graduate" or "some relevant experience" vs. "none"). Future work can test whether similar findings would emerge with these qualitative kinds of information.

Second, in Studies 2–3 of the present work, participants applied category boundaries consistently for different groups of targets. Category boundaries may be likely to remain consistent when people apply conventional definitions ("gain/loss," "donated/did not donate," "early/on time/late"), are given externally defined boundaries (e.g., a workplace defining performance cutoffs for a bonus or a school defining percentage cutoffs for grades), or when they have a wide range of experience and are less likely to shift their category boundaries in response to a couple new observations. In this vein, participants in Study 2 used externally defined boundaries (poor/moderate/excellent performance), and participants in Studies 3–4 intuitively applied conventional boundaries (gain/loss, some/none). However, fuzzy trace theory posits that categorical gist is itself context-dependent, shifting to reflect meaning within a given situation (Kühberger & Tanner, 2010; Reyna et al., 2021). In other settings, category boundaries might themselves shift to reflect a local context. For instance, a new manager might label a job candidate "borderline" in comparison to stronger candidates one day and might label a similar candidate "excellent" in comparison to weaker candidates the next day. Although social impressions could still reflect ordinal or categorical gist in these cases, categorical gist might not avoid reference-dependent decisions if category boundaries shift. The present work examines cases of stable category boundaries, thus revealing the underlying mental representations, but future work can test cases in which category boundaries shift. Nonetheless, the present work demonstrates that impressions can be encoded in different modes and that these modes shape social decisions.

Conclusions

The current work links social decision-making to fuzzy trace theory, suggesting that people encode social impressions as ordinal or categorical gists. Past work has tested the influence of semantic, reward-based, or episodic memory in social decision-making (Amodio, 2019; FeldmanHall et al., 2021; Hackel et al., 2020; Murty et al., 2016). The present work expands on these distinctions, suggesting that social impressions reflect fuzzy traces of social behavior encoded in two forms with distinct consequences for decision-making, highlighting how a multiplicity of mental representations can give rise to a multiplicity of decision-making styles in social choice.

References

Amodio, D. M. (2019). Social Cognition 2.0: An interactive memory systems account. *Trends in Cognitive Sciences*, *23*(1), 21–33. https://doi.org/10.1016/j.tics.2018.10.002

Barclay, P., & Willer, R. (2007). Partner choice creates competitive altruism in humans. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1610), 749–753. https://doi.org/10.1098/rspb.2006.0209

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). *Fitting linear mixed-effects models using lme4*. ArXiv. Preprint ArXiv:1406.5823. https://doi.org/10.48550/arXiv.1406.5823

Bendle, N. T. (2014). Reference dependence in political primaries. *Journal of Political Marketing*, *13*(4), 307–333. https://doi.org/10.1080/15377857.2012.721738

Biderman, N., & Shohamy, D. (2021). Memory and decision making interact to shape the value of unchosen options. *Nature Communications*, *12*(1), Article 4648. https://doi.org/10.1038/s41467-021-24907-x

Bless, H., & Schwarz, N. (2010). Mental construal and the emergence of assimilation and contrast effects: The inclusion/exclusion model. In M. P. Zanna (Eds.), *Advances in experimental social psychology* (Vol. 42, pp. 319–373). Elsevier.

Bornstein, A. M., Khaw, M. W., Shohamy, D., & Daw, N. D. (2017). Reminders of past choices bias decisions for reward in humans. *Nature Communications*, *8*(1), Article 15958. https://doi.org/10.1038/ncomms15958

Brainerd, C. J., & Reyna, V. F. (2001). Fuzzy-trace theory: Dual processes in memory, reasoning, and cognitive neuroscience. In H. W. Reese & R. Kail (Eds.), *Advances in child development and behavior* (Vol. 28, 41–100). Academic Press.

Brown, V. A. (2021). An introduction to linear mixed-effects modeling in R. *Advances in Methods and Practices in Psychological Science*, *4*(1), Article 2515245920960351. https://doi.org/10.1177/2515245920960351

Campion, M. A., Campion, J. E., & Hudson, J. P., Jr. (1994). Structured interviewing: A note on incremental validity and alternative question types. *Journal of Applied Psychology*, *79*(6), 998–1002. https://doi.org/10.1037/0021-9010.79.6.998

Campion, M. A., Pursell, E. D., & Brown, B. K. (1988). Structured interviewing: Raising the psychometric properties of the employment interview. *Personnel Psychology*, *41*(1), 25–42. https://doi.org/10.1111/j.1744-6570.1988.tb00630.x

Chang, L. W., & Cikara, M. (2018). Social decoys: Leveraging choice architecture to alter social preferences. *Journal of Personality and Social Psychology*, *115*(2), 206–223. https://doi.org/10.1037/pspa0000117

Corbin, J. C., Reyna, V. F., Weldon, R. B., & Brainerd, C. J. (2015). How reasoning, judgment, and decision making are colored by gist-based intuition: A fuzzy-trace theory approach. *Journal of Applied Research in*

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Memory and Cognition, *4*(4), 344–355. https://doi.org/10.1016/j.jarmac.2015.09.001

Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology*, *44*(1), 20–33. https://doi.org/10.1037/0022-3514.44.1.20

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, *34*(7), 571–582. https://doi.org/10.1037/0003-066X.34.7.571

Doll, B. B., Jacobs, W. J., Sanfey, A. G., & Frank, M. J. (2009). Instructional control of reinforcement learning: A behavioral and neurocomputational investigation. *Brain Research*, *1299*, 74–94. https://doi.org/10.1016/j.brainres.2009.07.007

Dotsch, R., Hassin, R. R., & Todorov, A. (2017). Statistical learning shapes face evaluation. *Nature Human Behaviour*, *1*(1), Article 0001. https://doi.org/10.1038/s41562-016-0001

Dunsmoor, J. E., & Murphy, G. L. (2014). Stimulus typicality determines how broadly fear is generalized. *Psychological Science*, *25*(9), 1816–1821. https://doi.org/10.1177/0956797614535401

Falvello, V., Vinson, M., Ferrari, C., & Todorov, A. (2015). The robustness of learning about the trustworthiness of other people. *Social Cognition*, *33*(5), 368–386. https://doi.org/10.1521/soco.2015.33.5.368

FeldmanHall, O., Montez, D. F., Phelps, E. A., Davachi, L., & Murty, V. P. (2021). Hippocampus guides adaptive learning during dynamic social interactions. *The Journal of Neuroscience*, *41*(6), 1340–1348. https://doi.org/10.1523/JNEUROSCI.0873-20.2020

Fiske, S. T. (1993). Social cognition and social perception. *Annual Review of Psychology*, *44*(1), 155–194. https://doi.org/10.1146/annurev.ps.44.020193.001103

Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Eds.), *Advances in experimental social psychology* (Vol. 23, pp. 1–74). Elsevier.

Foerde, K., Knowlton, B. J., & Poldrack, R. A. (2006). Modulation of competing memory systems by distraction. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(31), 11778–11783. https://doi.org/10.1073/pnas.0602659103

Hackel, L. M., & Mende-Siedlecki, P. (2023, June 5). *Two modes of social impressions and their effects on choice*. https://osf.io/xwvue/

Hackel, L. M., Mende-Siedlecki, P., & Amodio, D. M. (2020). Reinforcement learning in social interaction: The distinguishing role of trait inference. *Journal of Experimental Social Psychology*, *88*, Article 103948. https://doi.org/10.1016/j.jesp.2019.103948

Hackel, L. M., Mende-Siedlecki, P., Loken, S., & Amodio, D. M. (2022). Context-dependent learning in social interaction: Trait impressions support flexible social choices. *Journal of Personality and Social Psychology*, *123*(4), 655–675. https://doi.org/10.1037/pspa0000296

Hackel, L. M., & Zaki, J. (2018). Propagation of economic inequality through reciprocity and reputation. *Psychological Science*, *29*(4), 604–613. https://doi.org/10.1177/0956797617741720

Hester, N., & Gray, K. (2020). The moral psychology of raceless, genderless strangers. *Perspectives on Psychological Science*, *15*(2), 216–230. https://doi.org/10.1177/1745691619885840

Hewstone, M. (1994). Revision and change of stereotypic beliefs: In search of the elusive subtyping model. *European Review of Social Psychology*, *5*(1), 69–109. https://doi.org/10.1080/14792779543000020

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, *9*(1), 90–98. https://doi.org/10.1086/208899

Hunter, L. E., & Gershman, S. J. (2018). *Reference-dependent preferences arise from structure learning*. bioRxiv. https://doi.org/10.1101/252692.

Johnson, M. K., Kim, J. K., & Risse, G. (1985). Do alcoholic Korsakoff's syndrome patients acquire affective reactions?*Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*(1), 22–36. https://doi.org/10.1037/0278-7393.11.1.22

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, *39*(4), 341–350. https://doi.org/10.1037/0003-066X.39.4.341

Kalkstein, D. A., Hackel, L. M., & Trope, Y. (2020). Person-centered cognition: The presence of people in a visual scene promotes relational reasoning. *Journal of Experimental Social Psychology*, *90*, Article 104009. https://doi.org/10.1016/j.jesp.2020.104009

Kool, W., Gershman, S. J., & Cushman, F. A.. (2017). Cost-benefit arbitration between multiple reinforcement-learning systems. *Psychological Science*, *28*(9), 1321–1333. https://doi.org/10.1177/0956797617708288

Kühberger, A., & Tanner, C. (2010). Risky choice framing: Task versions and a comparison of prospect theory and fuzzy-trace theory. *Journal of Behavioral Decision Making*, *23*(3), 314–329. https://doi.org/10.1002/bdm.656

Mende-Siedlecki, P. (2018). Changing our minds: The neural bases of dynamic impression updating. *Current Opinion in Psychology*, *24*, 72–76. https://doi.org/10.1016/j.copsyc.2018.08.007

Mende-Siedlecki, P., Baron, S. G., & Todorov, A. (2013). Diagnostic value underlies asymmetric updating of impressions in the morality and ability domains. *The Journal of Neuroscience*, *33*(50), 19406–19415. https://doi.org/10.1523/JNEUROSCI.2334-13.2013

Mende-Siedlecki, P., Cai, Y., & Todorov, A. (2013). The neural dynamics of updating person impressions. *Social Cognitive and Affective Neuroscience*, *8*(6), 623–631. https://doi.org/10.1093/scan/nss040

Mills, B., Reyna, V. F., & Estrada, S. (2008). Explaining contradictory relations between risk perception and risk taking. *Psychological Science*, *19*(5), 429–433. https://doi.org/10.1111/j.1467-9280.2008.02104.x

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*(2), 61–64. https://doi.org/10.20982/tqmp.04.2.p061

Murty, V. P., FeldmanHall, O., Hunter, L. E., Phelps, E. A., & Davachi, L. (2016). Episodic memories predict adaptive value-based decision-making. *Journal of Experimental Psychology: General*, *145*(5), 548–558. https://doi.org/10.1037/xge0000158

Nisbett, R. E., & Miyamoto, Y. (2005). The influence of culture: Holistic versus analytic perception. *Trends in Cognitive Sciences*, *9*(10), 467–473. https://doi.org/10.1016/j.tics.2005.08.004

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review*, *108*(2), 291–310. https://doi.org/10.1037/0033-295X.108.2.291

Nowak, M. A., & Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature*, *393*(6685), 573–577. https://doi.org/10.1038/31225

Palminteri, S., Khamassi, M., Joffily, M., & Coricelli, G. (2015). Contextual modulation of value signals in reward and punishment learning. *Nature Communications*, *6*(1), Article 8096. https://doi.org/10.1038/ncomms9096

Palminteri, S., & Lebreton, M. (2021). Context-dependent outcome encoding in human reinforcement learning. *Current Opinion in Behavioral Sciences*, *41*, 144–151. https://doi.org/10.1016/j.cobeha.2021.06.006

Poldrack, R. A., Clark, J., Paré-Blagoev, E. J., Shohamy, D., Creso Moyano, J., Myers, C., & Gluck, M. A. (2001). Interactive memory systems in the human brain. *Nature*, *414*(6863), 546–550. https://doi.org/10.1038/35107080

R Core Team. (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Read, S. J. (1987). Constructing causal scenarios: A knowledge structure approach to causal reasoning. *Journal of Personality and Social Psychology*, *52*(2), 288–302. https://doi.org/10.1037/0022-3514.52.2.288

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Read, S. J., Jones, D. K., & Miller, L. C. (1990). Traits as goal-based categories: The importance of goals in the coherence of dispositional categories. *Journal of Personality and Social Psychology*, *58*(6), 1048–1061. https://doi.org/10.1037/0022-3514.58.6.1048

Reyna, V. F. (2012a). A new intuitionism: Meaning, memory, and development in fuzzy-trace theory. *Judgment and Decision Making*, *7*(3), 332–359. https://doi.org/10.1017/S1930297500002291

Reyna, V. F. (2012b). Risk perception and communication in vaccination decisions: A fuzzy-trace theory approach. *Vaccine*, *30*(25), 3790–3797. https://doi.org/10.1016/j.vaccine.2011.11.070

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences*, *7*(1), 1–75. https://doi.org/10.1016/1041-6080(95)90031-4

Reyna, V. F., Brainerd, C. J., Chen, Z., & Bookbinder, S. H. (2021). Explaining risky choices with judgments: Framing, the zero effect, and the contextual relativity of gist. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(7), 1037–1053. https://doi.org/10.1037/xlm0001016

Reyna, V. F., Chick, C. F., Corbin, J. C., & Hsia, A. N. (2014). Developmental reversals in risky decision making: Intelligence agents show larger decision biases than college students. *Psychological Science*, *25*(1), 76–84. https://doi.org/10.1177/0956797613497022

Richards, Z., & Hewstone, M. (2001). Subtyping and subgrouping: Processes for the prevention and promotion of stereotype change. *Personality and Social Psychology Review*, *5*(1), 52–73. https://doi.org/10.1207/S15327957PSPR0501_4

Rivers, S. E., Reyna, V. F., & Mills, B. (2008). Risk taking under the influence: A fuzzy-trace theory of emotion in adolescence. *Developmental Review*, *28*(1), 107–144. https://doi.org/10.1016/j.dr.2007.11.002

Sanfey, A. G. (2009). Expectations and social decision-making: Biasing effects of prior knowledge on ultimatum responses. *Mind & Society*, *8*(1), 93–107. https://doi.org/10.1007/s11299-009-0053-6

Satpute, A. B., Nook, E. C., Narayanan, S., Shu, J., Weber, J., & Ochsner, K. N. (2016). Emotions in "black and white" or shades of gray? How we think about emotion shapes our perception and neural representation of emotion. *Psychological Science*, *27*(11), 1428–1442. https://doi.org/10.1177/0956797616661555

Sharif, M. A., & Oppenheimer, D. M. (2016). The effect of relative encoding on memory-based judgments. *Psychological Science*, *27*(8), 1136–1145. https://doi.org/10.1177/0956797616651973

Sharif, M. A., & Oppenheimer, D. M. (2021). The effect of categories on relative encoding biases in memory-based judgments. *Organizational Behavior and Human Decision Processes*, *162*, 1–8. https://doi.org/10.1016/j.obhdp.2020.10.005

Simonsohn, U., & Gino, F. (2013). Daily horizons: Evidence of narrow bracketing in judgment from 10 years of MBA admissions interviews. *Psychological Science*, *24*(2), 219–224. https://doi.org/10.1177/0956797612459762

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research*, *16*(2), 158–174. https://doi.org/10.1086/209205

Sommerfeld, R. D., Krambeck, H.-J., Semmann, D., & Milinski, M. (2007). Gossip as an alternative for direct observation in games of indirect reciprocity. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(44), 17435–17440. https://doi.org/10.1073/pnas.0704598104

Spiers, H. J., Love, B. C., Le Pelley, M. E., Gibb, C. E., & Murphy, R. A. (2017). Anterior temporal lobe tracks the formation of prejudice. *Journal of Cognitive Neuroscience*, *29*(3), 530–544. https://doi.org/10.1162/jocn_a_01056

Todorov, A., & Olson, I. R. (2008). Robust learning of affective trait associations with faces when the hippocampus is damaged, but not when the amygdala and temporal pole are damaged. *Social Cognitive and Affective Neuroscience*, *3*(3), 195–203. https://doi.org/10.1093/scan/nsn013

Todorov, A., & Uleman, J. S. (2002). Spontaneous trait inferences are bound to actors' faces: Evidence from a false recognition paradigm. *Journal of Personality and Social Psychology*, *83*(5), 1051–1065. https://doi.org/10.1037/0022-3514.83.5.1051

Todorov, A., & Uleman, J. S. (2003). The efficiency of binding spontaneous trait inferences to actors' faces. *Journal of Experimental Social Psychology*, *39*(6), 549–562. https://doi.org/10.1016/S0022-1031(03)00059-3

Tranel, D., & Damasio, A. R. (1993). The covert learning of affective valence does not require structures in hippocampal system or amygdala. *Journal of Cognitive Neuroscience*, *5*(1), 79–88. https://doi.org/10.1162/jocn.1993.5.1.79

Wedekind, C., & Braithwaite, V. A. (2002). The long-term benefits of human generosity in indirect reciprocity. *Current Biology*, *12*(12), 1012–1015. https://doi.org/10.1016/S0960-9822(02)00890-4

Wedekind, C., & Milinski, M. (2000). Cooperation through image scoring in humans. *Science*, *288*(5467), 850–852. https://doi.org/10.1126/science.288.5467.850

Winter, L., & Uleman, J. S. (1984). When are social judgments made? Evidence for the spontaneousness of trait inferences. *Journal of Personality and Social Psychology*, *47*(2), 237–252. https://doi.org/10.1037/0022-3514.47.2.237

Witt, J. K. (2019). Graph construction: An empirical investigation on setting the range of the y-axis. *Meta-Psychology*, *3*, Article MP.2018.895. https://doi.org/10.5626/MP.2018.895

Wood, W. (2017). Habit in personality and social psychology. *Personality and Social Psychology Review*, *21*(4), 389–403. https://doi.org/10.1177/1088868317720362

Yoeli, E., Hoffman, M., Rand, D. G., & Nowak, M. A. (2013). Powering up with indirect reciprocity in a large-scale field experiment. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(supplement_2), 10424–10429. https://doi.org/10.1073/pnas.1301210110

Received June 5, 2022
Revision received March 20, 2023
Accepted March 31, 2023 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.