Judgment and Decision Making, Vol. 17, No. 2, March 2022, pp. 425–448

Violations of economic rationality due to irrelevant information during learning in decision from experience

Mikhail S. Spektor* Hannah Seidler†

Abstract

According to normative decision-making theories, the composition of a choice set should not affect people's preferences regarding the different options. This assumption contrasts with decades of research that have identified multiple situations in which this principle is violated, leading to context effects. Recently, research on context effects has been extended to the domain of experience-based choices, where it has been shown that forgone outcomes from irrelevant alternatives affect preferences — an accentuation effect. More specifically, it has been shown that an option presented in a situation in which its outcomes are salient across several trials is evaluated more positively than in a context in which its outcomes are less salient. In the present study, we investigated whether irrelevant information affects preferences as much as relevant information. In two experiments, individuals completed a learning task with partial feedback. We found that past outcomes from non-chosen options, which contain no relevant information at all, led to the same accentuation effect as did counterfactual outcomes that provided new and relevant information. However, if the information is entirely irrelevant (from options that could not have been chosen), individuals ignored it, thus ruling out a purely perceptual account of the accentuation effect. These results provide further support for the influence of salience on learning and highlight the necessity of mechanistic accounts in decision-making research.

Keywords: accentuation effect, context effects, decision making, decisions from experience, reinforcement learning

*Department of Economics and Business, Universitat Pompeu Fabra, and Barcelona School of Economics. Email: mikhail@spektor.ch. ORCID: 0000-0003-0652-1993.

†Department of Economics and Management, Karlsruhe Institute of Technology. Email: hannah.seidler@kit.edu. ORCID: 0000-0003-3055-3581.

The authors thank David Kellen for helpful comments and suggestions. Mikhail Spektor gratefully acknowledges financial support from the Spanish State Research Agency (AEI) through the Severo Ochoa Programme for Centres of Excellence in R&D (CEX2019–000915–S), the Spanish Ministry of Economic Affairs and Digital Transformation (MINECO) through the Juan de la Cierva fellowship (FJC2019–040970–I), the Spanish Ministry of Science and Innovation (MICINN; project PID2019–105249GB–I00), and the BBVA Foundation (project G999088Q). Data of the reported experiments and model codes are available at https://osf.io/s52z8/.

Copyright: © 2021. The authors license this article under the terms of the Creative Commons Attribution 3.0 License.

https://doi.org/10.1017/S1930297500009177 Published online by Cambridge University Press

1 Introduction

For centuries, scholars have been investigating whether humans make rational decisions (e.g., Bernoulli, 1738/1954), where "rationality" is defined as the conformity of choices to an axiomatic system of preferences (e.g., von Neumann & Morgenstern, 1947; Luce, 1959). Across different situations, human decision makers seem to violate every single principle of economic rationality (e.g., Müller-Trede et al., 2015; see Rieskamp et al., 2006, for an overview), which sparked the development of descriptively more adequate decision-making theories (e.g., Roe et al., 2001; see Busemeyer et al., 2019, for a recent review). Many of these theories aim to explain violations of *independence from irrelevant alternatives* (Luce, 1959), according to which preferences between any two options should be unaffected by addition or removal of other available options. Violations of this type are called *context effects*, such as the attraction effect (Huber et al., 1982), the compromise effect (Simonson, 1989), or the similarity effect (Tversky, 1972), and are among the most-studied phenomena in the decision-making literature (e.g., Trueblood et al., 2013).

Context effects are assumed to arise due to the multi-attribute nature of choice alternatives, with their attributes being evaluated not in isolation but in comparison to one another (e.g., Noguchi & Stewart, 2018; see Spektor et al., 2021 for a recent review). For example, a hiring officer might have to decide between job candidates that are equally qualified but differ in terms of their work experience and salary expectations. All theories that rely on a multi-attribute structure assume that attribute values are exactly known and accessible to the decision maker. However, in many situations, people have to infer the properties of choice alternatives from interactions with them. These decisions from experience have been shown to differ substantially from their description-based counterparts (Wulff et al., 2018). For example, when people are faced with a decision between two described lottery options with a discrete number of monetary rewards, they choose as if they are overweighting the probabilities of unlikely events (as proposed by prospect theory: Kahneman & Tversky, 1979). In contrast, when participants have to learn about the two options from experience, the opposite pattern occurs, a phenomenon that became known as the description-experience gap in risky choice (Hertwig & Erev, 2009). Research on experience-based choices bears a lot of potential to understand how contexts affect choices and the cognitive processes underlying them, as these types of decisions provide insights not only about how people make decisions but also about which representations of the options they obtain. For example, traditional context-effect research often relies on decisions between lottery options that are characterized by a single non-zero outcome (e.g., Herne, 1999; Tversky, 1972; Wedell, 1991; Soltani et al., 2012), where the outcomes and their corresponding probabilities span a two-dimensional attribute space. While it has been shown that classical context effects that rely on a multi-attribute structure of options can arise when people obtain such a representation from experience (Hadar et al., 2018), the required representation does not always arise (Hadar et al., 2018; Spektor et al., 2019; Ert & Lejarraga, 2018).

https://doi.org/10.1017/S1930297500009177 Published online by Cambridge University Press

TABLE 1: Illustration of outcome saliency underlying the accentuation effect

Stock	Month 1	Month 2	Month 3	Month 4	Month 5	Month 6
X	10	12	17	18	19	11
Y	16	16	12	12	12	19
Z	18	17	13	11	10	18

Despite the evidence that individuals often do not obtain such a representation in experience-based choices, their choices were nevertheless systematically influenced by the context (Spektor et al., 2019). This influence can best be described as follows: In a choice situation in which there are no clearly superior options, options whose rewards (outcomes) are particularly different (or distinct) from other rewards (repeatedly over trials) are chosen more often compared to a choice situation in which the very same outcomes are not as different from the other rewards — the *accentuation effect*. For example, consider the three stocks X, Y, and Z, whose values across six months are depicted in Table 1. The values of X and Y are negatively correlated, such that when X's value rises, the value of Y tends to decrease. If the value of Z is positively correlated with that of Y (and negatively with that of X), then the value of X is particularly distinct from the other two, and therefore perceived as more attractive.[1] For the accentuation effect to arise, it is not necessary for the decision maker to hold the kind of multi-dimensional attribute representation of the options that is available when options are described; Choice context affects choices by making certain rewards particularly distinct (e.g., from rewards that are negatively correlated over trials).

One of the main limitations of past research on context effects in experience-based choices has been the reliance on the *full-feedback paradigm* (Ert & Lejarraga, 2018; Spektor et al., 2019; but see Hadar et al., 2018, that used a different paradigm) in which individuals repeatedly make consequential choices between multiple options and receive feedback about the obtained and the forgone outcomes (i.e., from the options they did not choose). However, in many real-life situations, decision makers do not obtain counterfactual information about non-chosen options: for example, hiring officers might never find out how well the job candidates that were not hired would have performed had they joined the company. On the other hand, whenever forgone feedback is available, it is often highly relevant: if the hiring officers learn about the performance of the rejected job candidates, they might want to try hiring them at a later point in time. Therefore, the feedback stems from what we call *relevant alternatives*, as the decision maker is motivated to learn about the performance of the non-chosen alternatives (the rejected job candidates in the example) in order to try and choose (hire in the example) them in the future. However, processing the information

[1]Notably, the influence of distinctiveness is assumed to be sign-independent, so X would be chosen more often following a relatively bad-but-salient outcome than in a situation in which the same outcome is not as salient (e.g., when there are only two options available for choice).

https://doi.org/10.1017/S1930297500009177 Published online by Cambridge University Press

from the chosen and non-chosen options is cognitively taxing. Imagine the effort that hiring officers would need to exert trying to follow up on the performance of rejected job candidates at other companies. It is therefore expected that people will try to reduce the cognitive costs associated with processing forgone outcomes by engaging in a heuristic process, at the cost of potential loss of utility. For example, when tracking the performance of rejected job applicants, hiring officers might focus on those whose performance consistently differ from the other applicants (as in the case of the accentuation effect). Past research has demonstrated that merely paying attention to choice options increases the propensity to choose them (Cavanagh et al., 2014; Gluth et al., 2018), so hiring officers who focus their attention on specific candidates would be more likely to hire candidates with unique profiles in the future.

The goal of the present work is to shed light on the role of information relevance on the manifestation of the accentuation effect. Specifically, we investigated whether the accentuation effect also occurs when some of the information presented about non-chosen options is not informative, that is, it does not provide any new evidence in favor of or against choosing those non-chosen options. If this is the case, choices cannot be explained by a cost–benefit account, according to which context effects could arise as a by-product of the vast amount of information that has to be processed. Being aware of which pieces of information are informative, decision makers that maximize their rewards would ignore the ones that are not informative and focus on those that are. Foreshadowing our results, we found that individuals successfully ignored this information if it was not tied to the task they were solving. However, when information was task-related, it was processed similarly to how relevant information would be processed. The type of influence is consistent with a recently proposed learning model (Spektor et al., 2019) but not with other prominent theoretical accounts. For example, it cannot be explained by a higher decision weight for salient events (Bordalo et al., 2012) or by contextual value adaptation (Palminteri et al., 2015). Overall, our results suggest that context effects also occur in situations in which there is no need to allocate attentional resources between a large amount of information.

2 Experiment 1

In Experiment 1, we have investigated the accentuation effect in a setting in which individuals did not obtain counterfactual information about non-chosen options (i.e., information about what rewards the other options would have yielded had they been chosen). However, in contrast to similar tasks, we provided individuals with reminders about what they received from the non-chosen options when they chose them in the past. In contrast to a setting that provides counterfactual information, the additionally displayed reminders about the outcomes of non-chosen options do not contain any new information and are thus *irrelevant outcomes from relevant options* (relevant options since they are available for choice). Importantly, we argue that incorporating these outcomes into the preference formation of the

https://doi.org/10.1017/S1930297500009177 Published online by Cambridge University Press

chosen option reflects, in addition to violations of economic rationality, also an inefficient allocation of cognitive resources: Since the information provided about non-chosen options is not new, a rational decision maker should have incorporated this information into her evaluation of the option's value when that option was actually chosen; On every trial, she would fully focus on the outcome of the chosen option. Ignoring the irrelevant outcomes at the same time minimizes the amount of information that has to be processed and, therefore, the cognitive resources required. Moreover, this information is of unknown counterfactual validity to the individuals, so treating it as new information would bias their estimates.

2.1 Method

2.1.1 Participants and procedure

A total of 40 participants (29 female, 11 male, age 19–32, $M = 22.62$, $SD = 3.34$), mostly students of psychology at the University of Freiburg, with normal or corrected-to-normal vision, participated in the experiment. After giving informed consent, participants completed the experiment in individual cubicles. The procedure consisted of a demographic questionnaire, task instructions, a training block to assess learning performance, and two blocks of the experimental task (the order of which was counter-balanced across participants). In total, the experiment took approximately 45–60 min to complete and participants received the course-credit equivalent of an hour. Due to the hypothetical nature of the choices and the game-like framing of the task, we provided feedback about how many points they got in comparison to the other participants as motivation. We did not exclude any participants or trials. The behavioral data of both experiments and code for the computational models are available at https://osf.io/s52z8/.

2.1.2 Paradigm and materials

The paradigm used in the experiment was a heavily modified variant of the n-armed bandit problem (Sutton & Barto, 1998) with partial feedback. In an n-armed bandit problem, individuals repeatedly choose between (the same) n different options that provide monetary rewards according to their underlying outcome distributions. These outcome distribution are not known to the decision maker at the onset of the experiment. After each choice, they obtain a realization from the respective outcome distribution, thus learning which options yield the highest rewards through trial-and-error. In the present experiment, the outcome distributions of the options was comprised of the sum of three components: a systematic component, a constant (grand mean), and a noise component. The systematic component was based on three different events that occurred with certain probabilities, containing an option-specific outcome. Every time an event occurred, it yielded the same outcome. On top of this systematic component, a constant (or "grand mean") was added on every trial. This grand mean differed between participants and changed during the experiment multiple times, stemming from the value ranges 25–35, 35–45, and 45–55. Finally, a

https://doi.org/10.1017/S1930297500009177 Published online by Cambridge University Press

non-systematic noise component from a standard normal distribution was added on top of the other two components (see Figure 1A for the information display from the perspective of the participants). We developed the paradigm so that an isolated value representation is complicated whereas relating options' trial-by-trial outcomes to past outcomes of non-chosen options is comparatively easier.

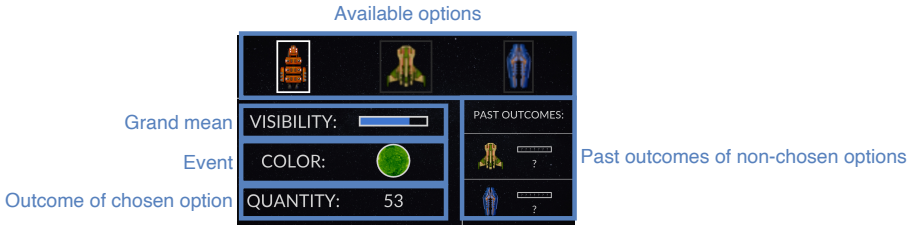
After a short inter-trial interval (400–600 ms), participants made a self-paced choice. The chosen option was highlighted for 900–1,100 ms and the non-chosen options were blurred, after which feedback about the current value of the grand mean, the current event, and the outcome of the chosen option on that trial was presented. This outcome was added to the participants' tally. Additionally, for every non-chosen option for which participants have encountered the same event (irrespective of the grand mean) at least once, they saw a reminder about what the grand mean and the outcome at the last observation with the respective event was. If the current event has not occurred when another option was chosen, then that option's reminder field remained blank. Importantly, the reminders did not contain any new information about the outcome distributions of the options. The feedback was presented for 4,000–4,500 ms, after which a new trial began.

The correlation between events and outcomes is an essential part of the experimental design. In a full-feedback setting, this structure is observable on a trial-by-trial basis. However, in a partial-feedback setting, non-chosen feedback is not provided, so it is not possible to establish a link between specific events and outcomes. For individuals to be able to relate events to outcomes, this information has to be conveyed explicitly. To do so and to increase task engagement, we framed it as an *extraterrestrial space mission* (similarly to Kool et al., 2016). Participants were told that a rare, valuable resource was found on extraterrestrial planets and that they were in charge of trying to retrieve as much of that valuable resource as possible. They had a selection of *probes* (representing the different options) available where each of the probes would try to retrieve as much of the rare resource as possible before returning to earth. Additionally, they were told that it was known that the amount retrieved depended on the color of a nearby star (representing the different events) and on the visibility on the planet (representing the grand mean). They were not told how each of these components related to each other. Figure 1B provides an illustration of a choice trial from the perspective of a participant. In this illustration, past feedback is available only for one of the two non-chosen options.

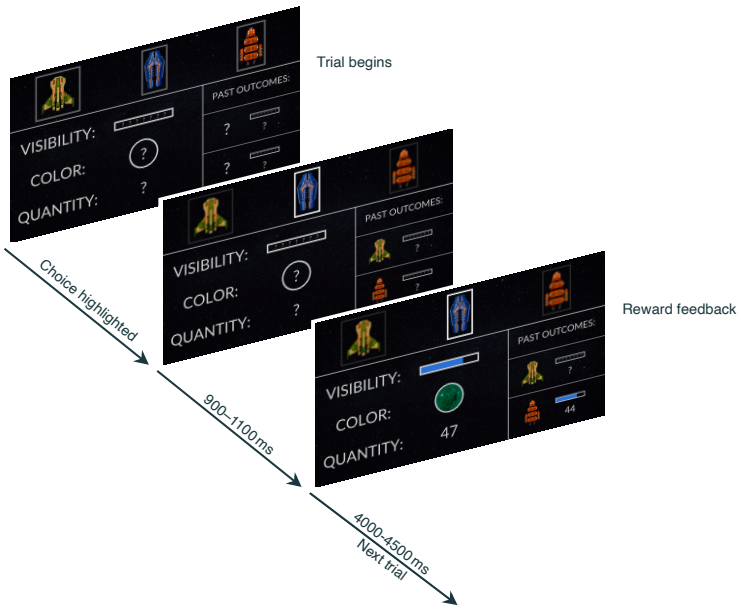
2.1.3 Design

A training block contained 40 trials in which individuals chose between a high-valued (HV) and a low-valued (LV) option and was used to assess general learning performance. The outcomes of the options depended on two events, $\mathcal{E}_1$ and $\mathcal{E}_2$, that occurred with probabilities $Pr(\mathcal{E}_1) = .6$ and $Pr(\mathcal{E}_2) = .4$. When $\mathcal{E}_1$ occurred, HV had an option-specific component of 33 and LV an option-specific component of -1. When $\mathcal{E}_2$ occurred, HV yielded –7 points and LV yielded –6 points. Therefore, the option-specific expected value (EV) of HV was

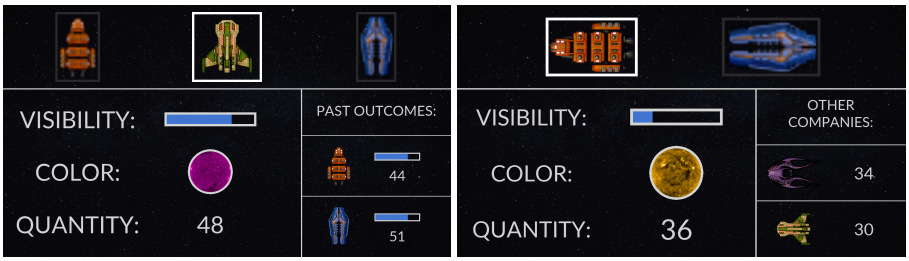
https://doi.org/10.1017/S1930297500009177 Published online by Cambridge University Press



(**A**) Choice screen



(**B**) Illustration of a choice trial



(**C**) Difference between Experiment 1 (left) and Experiment 2 (right)

FIGURE 1: Participants repeatedly chose between three options whose outcomes depended on *visibility*, *star color*, and an option-specific component. Initially, people did not have any information about how these interacted to form realized outcomes ("quantity extracted") and had to learn via trial-and-error. Additionally, participants received the information about the last quantity extracted with non-chosen options (if available). (**A**) Displays the relationship between the option components and the experimental task. (**B**) An example of a choice trial in Experiment 1. Here, the blue space ship was chosen. The orange space ship yielded 44 points the last time it was selected and the color of the star was green, while the yellow ship was not selected when the color of the star was green before. (**C**) In Experiment 2, there were no reminders of past outcomes but instead the outcomes of competitor companies. Visuals have been adapted for better readability. Spaceship images by MillionthVector are licensed under a Creative Commons Attribution 4.0 International License

https://doi.org/10.1017/S1930297500009177 Published online by Cambridge University Press

EV(HV) = 17 and EV(LV) = −3. The training block acquainted the participants with the task and was used to assess learning performance. The grand mean was initiated at the beginning of the training block and changed after 20 trials. Grand means comprised of a single draw from $\mathcal{U}(25, 35)$, $\mathcal{U}(35, 45)$, or $\mathcal{U}(45, 55)$, and the distribution from which they stemmed were drawn randomly without replacement.

The design of the experimental blocks was based on Spektor et al. (2019, Experiment 4: The Accentuation Effect), in which two options, B and C, were available for choice in two different choice sets of three options. The outcomes of the options depended on three events, $\mathcal{E}_1$, $\mathcal{E}_2$, and $\mathcal{E}_3$, that occurred with probabilities $Pr(\mathcal{E}_1) = .6$, $Pr(\mathcal{E}_2) = .3$, and $Pr(\mathcal{E}_3) = .1$. When $\mathcal{E}_1$ occurred, B had an option-specific component of –5.5 and C an option-specific component of –1.5. When $\mathcal{E}_2$ occurred, B yielded 6 points and C yielded 4 points. Finally, when $\mathcal{E}_3$ occurred, B resulted in a gain of 23 points and C in a gain of 5 points. In total, the options had the identical EV, with B being a riskier option (i.e., with a higher variance) than C.

The third option in each choice set of the experimental blocks, A or D, served as a *decoy* for C or B, respectively, therefore supposedly increasing that option's choice proportion relative to the other option. In other words, in the choice set $\mathcal{S}_1 = \{A, B, C\}$, C should be perceived as more attractive relative to B and in the choice set $\mathcal{S}_2 = \{D, B, C\}$, B should be perceived as more attractive relative to C. The outcome distributions of the options were constructed such that options A and B (C and D) yielded relatively similar outcomes on a trial-by-trial basis that are relatively dissimilar to option C (B): When event $\mathcal{E}_1$ ($\mathcal{E}_2$; $\mathcal{E}_3$) occurred, A resulted in a loss of 7 points (gain of 7 points; gain of 33 points) and D resulted in a gain of 2 points (0 points; 0 points).

To illustrate the effect of salience, consider the case of when event $\mathcal{E}_3$ occurs: In choice context $\mathcal{S}_1$, the options A, B, and C yield 33, 23, and 5 points, respectively. The most salient outcome (i.e., the outcome that is most dissimilar to the other outcomes) is 5 points of option C. In $\mathcal{S}_2$, options B, C, and D yield 23, 5, and 0 points, respectively. Here, the 23 points of option B are most salient, even though the outcomes of both B and C are identical across choice context (see Table 2 for a full description of the options and the choice sets they appear in).

Individuals completed the experimental blocks in a counter-balanced order and made 150 choices in each choice set. Grand means were drawn the same way as in the training block and changed twice within a block after 50 and 100 trials. The grand-mean changes were included to (1) encourage continuous learning in the task, (2) conceal that two of the options are identical across choice sets, and (3) to invalidate the past outcomes as counterfactuals (across grand means). In all cases, event occurrences were pseudo-randomly generated to be representative every 10 trials within an option. To avoid perceptual saliency effects, outcomes were truncated at 10 and 99. Associations between stimuli and the events, options, and choice sets they represent were randomized across participants.

Learning performance was quantified using two different dependent variables: raw

https://doi.org/10.1017/S1930297500009177 Published online by Cambridge University Press

TABLE 2: Choice set composition and the respective option-specific components

Set(s)	Option	Event (probability) $\mathcal{E}_1$ (.6)	$\mathcal{E}_2$ (.3)	$\mathcal{E}_3$ (.1)	EV	SD
$\mathcal{S}_1$	A	–7	7	33	1.2	12.31
$\mathcal{S}_1, \mathcal{S}_2$	B	-5.5	6	23	0.8	9.01
	C	–1.5	4	5	0.8	2.83
$\mathcal{S}_2$	D	2	0	0	1.2	0.98

Note. Option-specific components of the options were tied to the occurrence of events. See Design for details.

accuracy and corrected accuracy. Raw accuracy reflects the proportion of HV choices in the training block $Pr(\text{HV})$. Values of $Pr(\text{HV}) > .5$ reflect that individuals were able to learn that HV has a higher EV than LV. However, due to the random nature of observed outcomes, LV can have yielded better outcomes than HV for a limited number of observations. Corrected accuracy controls for the influence of sampling error by computing the running mean (i.e., the mean of all outcomes previously observed) within each option. A "correct" response is therefore choosing the option that has the higher running mean.

The manifestation of context effects was quantified using the *relative choice share of the target* (RST; Berkowitsch et al., 2014), where *target* is the option whose attractiveness is supposed to increase according to the accentuation effect: $\text{RST} = \frac{Pr(T)}{Pr(T)+Pr(C)}$ where $Pr(T)$ is the proportion of target choices (i.e., C in choice set $\mathcal{S}_1$ and B in choice set $\mathcal{S}_2$, respectively) and $Pr(C)$ is the proportion of competitor choices (i.e., B in choice set $\mathcal{S}_1$ and C in choice set $\mathcal{S}_2$, respectively). RST values range from 0 (competitor is always chosen) to 1 (target is always chosen), where RST = .50 indicates an absence of a context effect. RST > .50 indicates the presence of an accentuation effect. By using the RST as a dependent measure, we automatically control for individual prior preferences for low- or high-variance options (i.e., safe or risky options, respectively).

We also checked for violations of a less restrictive variant of the independence axiom, *weak independence from irrelevant alternatives* (see Rieskamp et al., 2006). This weaker axiom is violated if significantly more people prefer C over B in $\mathcal{S}_1$ while simultaneously preferring B over C in $\mathcal{S}_2$ than the other way round. In contrast to the stronger axiom, it does not restrict the choice proportions to be exactly equal but only requires the ordering of choices to remain stable across contexts.

2.1.4 Computational modeling

To assess the influence of irrelevant outcomes on a trial-by-trial basis, we analyzed the data in two different ways: First, we assessed whether the probability of repeating the same

https://doi.org/10.1017/S1930297500009177 Published online by Cambridge University Press

choice can be predicted by the obtained reward and the chosen option's salience using a logistic regression. Second, in order to obtain a mechanistic understanding of the cognitive processes underlying learning and decision making in the task, we used a formal modeling approach.

Within the context of the regression analysis, we relied on two predictors:

- The difference between the obtained reward and the running mean of the chosen option as the first predictor. In other words, it is positive if the obtained outcome is above the average of that option's previous outcomes and negative if it is below average and it reflects the degree to which individuals are sensitive to outcomes. Essentially, it corresponds to a reward-prediction error which is the standard learning signal in the literature (e.g., Schultz et al., 1997).
- The chosen option's salience, which is the centered standardized mean pairwise Euclidian distance between all outcomes presented on the screen (i.e., the obtained outcome and the past outcomes of non-chosen options, if available). The standardization achieves that all saliences are between 0 and 1, such that the predictor variable reflects the deviation from an "average" salience in every trial. This measure of salience is one that has relatively few assumptions and satisfies the following properties: (1) If only one or two outcomes are observed, then salience has no influence, (2) values above 0 (below 0) indicate that the chosen option's outcome is more (less) salient than the past outcomes from non-chosen options, and (3) a value of 0 indicates that the chosen option's outcome has an average salience.

Regression weights of the reward predictor above 0 reflect that individuals are sensitive to rewards: If the obtained outcome is better than what they expected, they are more likely to choose the same option again. Regression weights of the chosen option's salience above 0 reflect that individuals compare the chosen option's outcome with past outcomes of non-chosen options in line with the similarity mechanism: If an option's outcome is particularly salient on a given trial, then individuals are more likely to choose it again (compared to a situation in which the outcome is not as salient).

For the formal-modeling analysis, we fit a total of three nested reinforcement-learning models, two commonly used reinforcement-learning models that do not assume an influence of irrelevant outcomes and rigorously compared it to the *accentuation of differences model* that assumes such an influence (see Spektor et al., 2019). The first and most simple model is the *basic reinforcement learning model*. It keeps track of the subjective expectation $Q_{i,t}$ of option i on trial t and updates it using the reward-prediction error:

$$Q_{i,t+1} = Q_{i,t} + \alpha\big(R_{i,t} - Q_{i,t}\big), \tag{1}$$

where $R_{i,t}$ is the reward obtained on trial t. The only parameter of the basic reinforcement learning model is learning rate α, ranging from 0 to 1 and governing the degree to which individuals adapt to recent rewards.

https://doi.org/10.1017/S1930297500009177 Published online by Cambridge University Press

The second model is the *marginal utility function model* that nests the basic reinforcement learning model. In contrast to the latter, the former assumes a power function that maps observed rewards onto subjective utilities:

$$Q_{i,t+1} = Q_{i,t} + \alpha\left(R_{i,t}^{\gamma} - Q_{i,t}\right). \tag{2}$$

In the case of $\gamma = 1$, the marginal utility function model reduces to the basic reinforcement learning model. Values of γ above 1 (between 0 and 1) represent risk-seeking (risk-averse) behavior.

Finally, the accentuation of differences model assumes that subjective utilities are not evaluated in isolation but rather that particularly salient rewards receive more attention and are in turn perceived as more attractive (relative to less salient rewards). This intuition is implemented in form of an inhibitory similarity mechanism, which conceptually corresponds to inverse saliency. The same intuition applies: The more similar (i.e., closer on the number line) an option's outcome is to the other options' outcomes, the less attractive it becomes. Formally, $R_{i,t}^{\gamma}$ in Equation 2 is replaced by

$$R_{i,t}^{\gamma} - \eta \times Z \times \overline{R_{i,t}^{\gamma}}, \tag{3}$$

where Z is the average negative exponential distance between the respective option's perceived reward and the other perceived rewards,

$$Z = \frac{\sum_{j=1}^{J} \mathrm{e}^{-\psi \times |R_{i,t}^{\gamma} - R_{j,t}^{\gamma}|}}{J}, \tag{4}$$

and $\overline{R_{i,t}^{\gamma}}$ is the average perceived reward of all outcomes that scales Z up from a (0, 1) to a standardized scale (i.e., average perceived reward). The set J contains the last-seen rewards of non-chosen options for the same event that occurred.

The core parameter that determines the degree to which individuals take the similarity mechanism into account is η; $\eta = 0$ reflects an agent that ignores saliency, $\eta > 0$ is the standard case in which saliency increases an option's attractiveness, and $\eta < 0$ reflects a situation in which saliency reduces an option's attractiveness. Additionally, the scaling parameter ψ determines the sensitivity to the numerical distance between outcomes (see Spektor et al., 2019, for additional details and validations in a full-feedback setting).

To transform subjective expectations into choice probabilities, we used a soft-max choice rule with choice-sensitivity parameter θ:

$$Pr(i,t) = \frac{\mathrm{e}^{\theta Q_{i,t}}}{\sum \mathrm{e}^{\theta Q_{j,t}}}. \tag{5}$$

We fit each model within a hierarchical Bayesian framework (see Gelman et al., 2013, for an introduction) and compared the models using the *leave-one-out information criterion*

https://doi.org/10.1017/S1930297500009177 Published online by Cambridge University Press

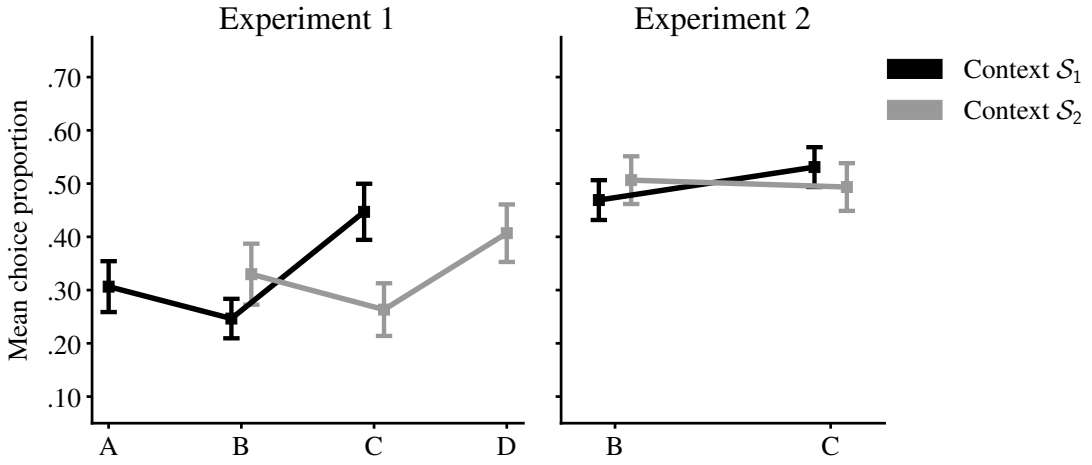


FIGURE 2: Mean choice proportions of the options in the two experiments and contexts. Context $\mathcal{S}_1$ comprised of options {A, B, C} and context $\mathcal{S}_2$ comprised of options {B, C, D}. Options' outcomes were tied to the occurrence of events, where options A and D served as decoys for options C and B, respectively. In Experiment 2, outcomes of options A and D (in the respective context) were presented as outcomes from non-available options. Error bars indicate the 95% CI of the mean.

(LOOIC; Vehtari et al., 2017). The LOOIC quantifies how well a model can explain the data and penalizes models that are complex to avoid overfitting. It does so by computing the effective number of parameters that does not depend directly on the number of free parameters but rather on how parameter values affect model predictions (see Vehtari et al., 2017, for details). Lower LOOICs reflect better penalized-for-complexity model fits. Models have been specified using weakly informative priors.

2.2 Results

2.2.1 Behavioral analyses

First, we checked whether participants chose the higher-valued option HV more often than they chose the lower-valued option LV in the training block. A one-sample t test on raw accuracy against .5 confirmed that HV was chosen in more than half of the cases (M = .65, SD = .10; $t(39)$ = 9.45, $p < .001$, d = 1.49, 95% CI [1.04, 1.94]). A one-sample t test on corrected accuracy led to the same conclusion (M = .66, SD = .11; $t(39)$ = 9.39, $p < .001$, d = 1.48, 95% CI [1.03, 1.93]).

Second, we checked whether individuals' choices violated the independence axiom (and, therefore, economic rationality). A violation of independence would be reflected in a significant change of the relative preference of options B and C across the two choice sets $\mathcal{S}_1$ and $\mathcal{S}_2$ (see Figure 2, left panel, for mean choice proportions in the two choice sets and Figure 3, top row, for aggregated choice proportions in bins of 10 trials). Behavior in line with the independence axiom would result in RSTs = .5 and the presence of an accentuation

https://doi.org/10.1017/S1930297500009177 Published online by Cambridge University Press

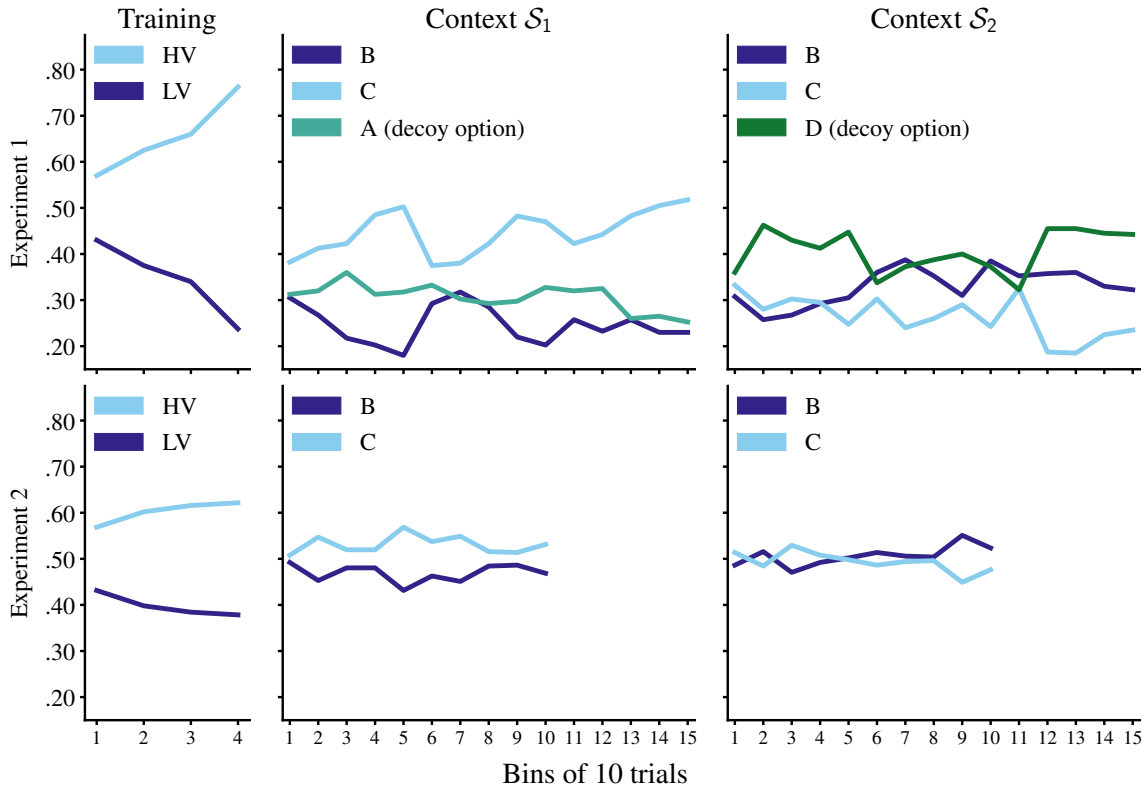


FIGURE 3: Aggregated choice proportions of the options in the two experiments and all choice sets, in bins of 10 trials. The training set comprised of a high-valued option HV and a lower-valued option LV and was used to assess learning performance. All options' outcomes were tied to the occurrence of events (star colors), where options A and D served as decoys for options C and B, respectively. In Experiment 2, outcomes of options A and D (in the respective context) were presented as outcomes from non-available options.

effect would result in RSTs > .5. A one-sample t test on RSTs against .5 confirmed the presence of a substantial accentuation effect ($M_{\text{RST}} = .59$, $SD_{\text{RST}} = .14$; $t(39) = 4.05$, $p < .001$, $d = 0.64$, 95% CI [0.30, 0.98]), where people chose the target option on average almost 50% more often than the competitor.

We followed up this analysis with a test for violations of "weak" independence from irrelevant alternatives which, if violated, contradicts more fundamental principles of economic rationality; This principle states that while relative choice proportions (e.g., the RST) can vary across contexts, modal choices should not. In contrast to this notion, 22 out of 40 individuals (55%) chose C more often than B in choice set $\mathcal{S}_1$ but B more often than C in choice set $\mathcal{S}_2$. In contrast, only 4 out of 40 individuals (10%) had the opposite pattern, which is the control condition to rule out random fluctuation. A 2×2 χ^2 contingency test confirmed the difference in preference-shift proportions ($\chi^2(1) = 16.47$, $p < .001$).

https://doi.org/10.1017/S1930297500009177 Published online by Cambridge University Press

2.2.2 Computational modeling

We investigated whether the observed violation of economic rationality can be explained by an attentional salience mechanism. According to such a mechanism, particularly salient outcomes of the chosen option increase and particularly non-salient outcomes of the chosen option decrease its attractiveness. For each individual, we performed logistic regressions on the probability of choosing the same option again with the reward-prediction error and the chosen option's outcome salience as predictors. A one-sample t test on the reward-prediction-error regression weight ($M = 0.02$, $SD = 0.04$) confirmed that individuals were sensitive to rewards ($t(39) = 3.81$, $p < .001$, $d = 0.60$, 95% CI [0.26, 0.94]). Crucially, the chosen option's outcome salience also incrementally predicted choice-repetition probability, as confirmed by a one-sample t test on the respective regression weight ($M = 3.24$, $SD = 4.58$; $t(39) = 4.47$, $p < .001$, $d = 0.71$, 95% CI [0.36, 1.05]).

Finally, we have compared the accentuation of differences model, a computational model that formalizes the cognitive processes supposedly underlying learning and decision making in the task, to two alternative models — nested within the accentuation of differences model—in their ability to explain the trial-by-trial choices of individuals (see Table 3 for results of the model comparison). The critical difference between the accentuation of differences model and the other reinforcement-learning models is that the accentuation of differences model assumes a mechanism that leads to outcome-salience dependent valuation of options. In line with the model-free analysis, our model comparison revealed that the accentuation of differences model provides a better account of the data (LOOIC = 22,389, $SE = 112.69$) than the utility-function model, the better of the other two models (LOOIC = 22,802, $SE = 106.84$), ΔLOOIC = 413 ($SE = 44.36$), resulting in a standardized effect size of 9.31σ, which means that the predictions of the better model are 9.31 standard errors away from those of the worse model (see Vehtari et al., 2017, for details). An additional model comparison with a basic reinforcement learning model that updates the expectations of non-chosen options using the reminders of past outcomes provided the worst account of the data (LOOIC = 23,650, $SE = 115.06$), a performance even below the chance level of 26,367 (after accounting for model complexity), ruling out that people simply confused the presented reminders with actual forgone outcomes or valid counterfactual outcomes.

The obtained group-level parameter estimates of the accentuation of differences model shed light on the cognitive processes at work. With a mean learning rate of $\alpha = .04$, individuals have a rather long time window of integration. The mean curvature of the utility function reflects a moderate degree of risk aversion, $\gamma = 0.61$, and the parameter that determines the degree of similarity-based inhibition is positive, $\eta = 0.36$ (however, its highest-density interval overlaps with 0, suggesting a substantial degree of individual differences). See Table 4 for a summary of the group-level posterior.

https://doi.org/10.1017/S1930297500009177 Published online by Cambridge University Press

TABLE 3: Information Criteria for Each of the Models in Both Experiments

Experiment	Model	LOOIC	p_{LOOIC}	SE_{LOOIC}
1	Basic reinforcement learning model	22,901	89	106.82
	Marginal utility function model	22,802	96	106.84
	Accentuation of differences model	22,389	128	112.69
2	Basic reinforcement learning model	13,401	48	52.78
	Marginal utility function model	13,406	49	52.49
	Accentuation of differences model	13,332	104	56.52

Note. LOOIC = Leave-one-out information criterion (Vehtari et al., 2017). p_{LOOIC} = Effective number of parameters. SE_{LOOIC} = Standard error of the LOOIC. All measures are reported on the deviance scale.

TABLE 4: Posterior Distributions of the Accentuation of Differences Models' Group-Level Mean Parameters

			Posterior percentile	
	Parameter	M	2.5%	97.5%
Experiment 1	θ	0.37	0.20	0.64
	α	.04	.02	.07
	γ	0.61	0.46	0.81
	ψ	1.27	0.44	3.71
	η	0.36	-0.46	1.19
Experiment 2	θ	0.08	0.03	0.20
	α	.02	.01	.04
	γ	1.00	0.77	1.27
	ψ	1.64	0.48	5.82
	η	0.14	-0.50	0.90

Note. See Computational modeling of Experiment 1 for a detailed description of the model and its respective parameters.

2.3 Discussion

The accentuation effect is a context effect that emerges from the trial-by-trial salience of outcomes in a learning setting (Spektor et al., 2019). Experiment 1 investigated this effect in an experience-based setting with partial feedback and reminders of past outcomes from non-chosen options. We found that in such a setting, individuals showed an accentuation effect of considerable size. Analyses based on a varying degree of assumptions, ranging from a

https://doi.org/10.1017/S1930297500009177 Published online by Cambridge University Press

logistic regression to a full-fledged model comparison, confirmed the assumed mechanism underlying the accentuation effect, namely that dissimilar outcomes are perceived as more attractive.

Previous comparable studies (Ert & Lejarraga, 2018; Spektor et al., 2019) used a full-feedback paradigm in which individuals obtained counterfactual feedback about the rewards of the non-chosen options (i.e., the reward they would have earned had they chosen them). Compared to partial feedback, full feedback substantially facilitates the task as individuals obtain information about forgone outcomes free-of-cost. In contrast, individuals in the partial-feedback situation have to trade off forgone rewards from not-choosing the option they think is the best with the possibility of finding an option that is even better — an *exploration–exploitation dilemma* (e.g., Navarro et al., 2016). More importantly, individuals cannot compare the outcomes of the options with each other, a necessary condition for certain rewards to be particularly salient and, therefore, for the accentuation effect to arise.

To compensate for this property of the partial-feedback paradigm, Experiment 1 substantially deviated from the full-feedback paradigm by not only leaving out the forgone feedback but also by providing individuals with information about the structure of the environment (information about the star color and visibility referring to the grand-mean component and the event, respectively); Information that individuals in the full-feedback paradigm did not obtain. Most notably, individuals also saw the outcomes they have gotten from non-chosen options in the past. While we found no evidence that individuals treated these outcomes as actual outcomes or valid counterfactuals, it is possible that individuals still perceived these reminders to be informative of what they would have gotten had they chosen the respective option, especially since these options were relevant for them.

Given the design of the experiment, whenever individuals have observed the outcomes of both non-chosen options in the same event and the same grand-mean component, the outcomes were in fact not too far away from being valid counterfactuals. However, this was somewhat rarely the case, especially for the less frequent events; Only in 81% of the trials did participants actually obtain the reminders from *both* of the non-chosen options, and only in 77% of these cases (63% of the total trials) was the information from the same grand-mean component. Put differently: In 37% of all trials, paying any attention at all to the past outcomes would introduce a non-negligible bias into any estimate based on them. Only a sophisticated understanding of how the different components relate to each other could correct for this bias. It is unlikely that participants obtain such an understanding and correct for the bias in order to draw counterfactual conclusions about the non-chosen options. We therefore argue that the past outcomes from non-chosen options were, indeed, entirely "irrelevant". Experiment 2 aimed to explore how grave this misperception is: Do individuals blindly react to fully irrelevant information, much like with *anchoring* (Tversky & Kahneman, 1974) or does it only occur when the information is factually irrelevant but stems from relevant options?

https://doi.org/10.1017/S1930297500009177 Published online by Cambridge University Press

3 Experiment 2

The first experiment demonstrated how normatively irrelevant feedback can lead to context-dependent learning in a partial-feedback setting. However, the irrelevant feedback came from relevant options, so individuals might process this information as if it was relevant. The goal of the second experiment was to investigate the effect of irrelevant information from irrelevant options. More precisely, it was to test whether the influence of irrelevant information is goal independent and occurs in any situation in which it is available (much like a purely perceptual phenomenon) or whether individuals only process such information if it is goal relevant.

In order to address this question, Experiment 2 flipped around the logic of Experiment 1: Instead of providing de facto irrelevant information from relevant options, we provided individuals with valid counterfactual information from irrelevant information sources. If the accentuation effect is mainly a perceptual phenomenon based on numerical saliency, it should arise in Experiment 2 as well; After all, the visual presentation is mostly identical in both experiments.

3.1 Method

The experiment was a modified variant of Experiment 1, with the following differences. A total of 51 participants (30 female, 21 male, age 18–22, $M = 20.12$, $SD = 1.00$), mostly students with different majors from the Universitat Pompeu Fabra, Barcelona, took part in the experiment. The experiment took approximately 30–40 min to complete and participants received a show-up fee of 5 Euro and a choice-dependent bonus of up to 4 Euro. We did not exclude any participants or trials.

The experimental part of the task was a modified version of the one from the first experiment. Instead of three options in the choice set (and 150 decisions), participants always chose between two options, namely B and C, for 100 trials in each context. Participants received feedback about the outcome of the chosen option, the current event, and the grand mean, much like in Experiment 1. However, the "irrelevant feedback" they received this time stemmed from two options that were explicitly labeled as unavailable to them. The outcomes of the non-available options corresponded to the counterfactual forgone outcome of the non-chosen option and hypothetical counterfactual outcomes of option A (context $\mathcal{S}_1$) and option D (context $\mathcal{S}_2$). For example, if option B was chosen, the outcomes of the non-available options were the forgone outcome of option C and the outcome of option A (or D in the other context). On each trial, these values were randomly mapped to the two non-available space ships so individuals could not learn that the outcome of one of the non-available ships in fact corresponded to the other available option. Figure 1C illustrates the difference between the two experiments.

https://doi.org/10.1017/S1930297500009177 Published online by Cambridge University Press

3.2 Results

We confirmed that participants chose the higher-valued option HV more often than the lower-valued option LV in the training block. A one-sample t test on raw accuracy against .5 confirmed that HV was chosen in more than half of the cases (M = .60, SD = .11; $t(50)$ = 6.40, $p < .001$, d = 0.90, 95% CI [0.57, 1.22]). A one-sample t test on corrected accuracy led to the same conclusion (M = .61, SD = .12; $t(50)$ = 6.50, $p < .001$, d = 0.91, 95% CI [0.58, 1.23]).

In contrast to Experiment 1, we did not find a significant change in the choice proportions, as reflected in a one-sample t test on RSTs against .5 (M_{RST} = .52, SD_{RST} = .10; $t(50)$ = 1.32, p = .19, d = 0.18, 95% CI [-0.09, 0.46]). See Figure 2, right panel, for mean choice proportions in the two choice sets and Figure 3, bottom row, for aggregated choice proportions in bins of 10 trials. In line with this, the test for violations of the weak version of the independence principle showed that 13 out of 51 participants (25%) had chosen C more often than B in $\mathcal{S}_1$ and at the same time B more often than C in $\mathcal{S}_2$, with 7 individuals (14%) showing the opposite pattern. The difference was not significant, as confirmed by a 2×2 χ^2 contingency test ($\chi^2(1) = 1.55$, p = .21).

In line with the main behavioral results, the logistic regression revealed a significant influence of the reward-prediction error on the probability to repeat the previous choice (i.e., reward sensitivity; M = 0.01, SD = 0.03; $t(50)$ = 2.24, p = .03, d = 0.31, 95% CI [0.03, 0.59]), but no influence of the chosen option's outcome salience (M = 0.65, SD = 3.99; $t(50)$ = 1.16, p = .25, d = 0.16, 95% CI [-0.12, 0.44]). The high individual variability in the salience weighting was reflected in the model comparison (see Table 3), where the accentuation of differences model provided the best account of the data (LOOIC = 13,332, SE = 56.52), but only with a small margin to the second-best model , with ΔLOOIC = 70 (SE = 22.99) and a standardized effect size of 3.04σ (see Table 4 for a summary of the group-level posterior of the accentuation of differences model).

Given that participants were not randomly allocated to the two experiments, a direct comparison between them is not possible. Nevertheless, descriptively, the effect sizes obtained in Experiment 2 are consistently lower than in Experiment 1, suggesting a consistently lower degree of context dependency.

3.3 Discussion

Experiment 2 aimed to distinguish whether the accentuation effect is a purely perceptual phenomenon or whether it is related to goal-related processes. To do so, we have flipped around the logic of the first experiment by providing individuals with new information that stemmed from irrelevant alternatives. In this setting we found no evidence of a pronounced accentuation effect, as all analyses agreed that the accentuation effect does not arise in this setting.

https://doi.org/10.1017/S1930297500009177 Published online by Cambridge University Press

4 General Discussion

The present work investigated whether individuals form preferences in learning tasks independently of irrelevant outcomes. In contrast to notions of economic rationality, we found that preferences shift depending on the choice context and that these preference shifts are driven by irrelevant outcomes, but only if these irrelevant outcomes stem from relevant options. In such a situation, our results support the recently established notion that particularly salient outcomes on a trial-by-trial basis increase that option's perceived attractiveness.

4.1 A model of the experiment

So far, the accentuation effect has been investigated in a full-feedback paradigm only (Spektor et al., 2019). In this paradigm, the psychological process supposedly underlying it is a rather straight-forward process: It is easy to compare outcomes with one another on a trial-by-trial basis and discount options whose outcomes are similar. Not only does the exploration–exploitation dilemma make the partial-feedback setting considerably more complex, but it is also not possible to directly compare outcomes with each other. This increased complexity has been shown to result in slower learning (Yechiam & Busemeyer, 2005), lower choice accuracy (Rakow et al., 2015; Yechiam & Rakow, 2012; Palminteri et al., 2015), and a higher impact of surprising outcomes (Plonsky & Erev, 2017). In order to facilitate learning in the task and isolate the expected effect of outcome saliency on choices, we have provided individuals with some structural information about the task, information that is typically not provided in experience-based paradigms.

In such a setting, we were able to show that the mere presence of irrelevant outcomes affects preference formation. However, this was only the case when the irrelevant information came from relevant choice alternatives. Whenever the information came from supposedly entirely irrelevant sources, individuals successfully ignored it. This indicates that the irrelevant information of relevant options is interpreted as relevant information and is considered in the decision-making process. These results speak against the notion that the accentuation effect is a perceptual phenomenon that is insensitive to the relevance of sources. Nevertheless, the use of irrelevant information from relevant alternatives poses a violation of normative principles. Any kind of reprocessing of past outcomes as relevant information would lead to a biased estimate of the perceived value of the currently chosen option, the non-chosen option, or both. Additionally, a comparison of the past outcome with the outcome of the currently chosen option (in line with the assumed mechanism underlying the accentuation effect) would violate the independence principle and lead to context effects. Even less rigid extensions of economic rationality, such as those that assume a cost–benefit analysis of information acquisition, would predict that the irrelevant information should be ignored: irrespective of how it is processed, not processing it at all is less effort than even the most heuristic kind of processing.

https://doi.org/10.1017/S1930297500009177 Published online by Cambridge University Press

It is noteworthy that the experimental setting might have suggested to the individuals that the information that is presented to them is somehow relevant and that they should use it, despite explicit instructions about the factual irrelevance thereof. In this case, individuals would be solving a different problem than what the experimenters expect them to solve (e.g., Szollosi & Newell, 2020; Kellen, 2019). We see no strong reason to suspect that we are dealing with such a situation: A model that explicitly treats past outcome reminders as if it was valid information cannot account for the behavior observed in Experiment 1. Moreover, while Experiment 2 provided more information for individuals to make use of (even though they were not aware of that fact), it did not significantly affect participants' behavior. Finally, even if individuals felt they had to use the information presented to them somehow, it is doubtful the information carries any suggestion in line with the mechanism giving rise to the accentuation effect. In sum, the behavior observed in the present experiments is unlikely to occur due to experimental demand effects.

4.2 Broader relevance of accentuation effects

While the present study was designed to elicit the strongest accentuation effect possible, this specific experimental setup is not necessary for accentuation effects to arise. Situations in which individuals get partial feedback along with reminders of past choices (e.g., when online shops remind their customers of past purchases or when streaming services provide reminders of already-watched movies) are quite common in everyday life, and accentuation effects are expected to occur in these situations as well. Importantly, these reminders are often of questionable informational relevance. The present study sheds light on how individuals are susceptible to the influence of particularly distinct outcomes in such situations and the role of informational relevance.

Traditionally, context-effects research has relied on choice options that are each described on two attribute dimensions (Tversky, 1972; Huber et al., 1982; Simonson, 1989; Trueblood et al., 2013). The accentuation effect breaks with this tradition by not being defined in terms of an interaction between attribute dimensions but by the trial-by-trial reward dynamics. Although both types of context effects constitute violations of the independence axiom, their qualitative differences raise the question whether the effects belong to the same or to distinct categories of context effects. So far, the two types of context effects do not seem to arise within the same setting. Future research should clarify the degree to which a separate treatment is necessary or not.

4.3 Alternative models, possible explanations, and conclusion

The present study used a similarity mechanism with an underlying reinforcement-learning mechanism to interpret participants' behavior. Here, we will discuss whether the observed behavior is compatible with alternative theoretical approaches, even though we are not

https://doi.org/10.1017/S1930297500009177 Published online by Cambridge University Press

aware of any alternative model that would be able to account for the observed choices without additional assumptions and adaptations.

The semantically most closely related model is surely salience theory (Bordalo et al., 2012), according to which particularly salient outcome states receive a higher decision weight, where salience is essentially the range of outcomes. As a theory of decisions under risk (that assumes perfect knowledge of the options' outcome distributions), two possible modifications to the setting of repeated choices come to mind. First, individuals might learn the reward contingencies explicitly (as displayed in Table 2). Second, individuals might use the trial-by-trial salience to determine the degree to which they update reward expectations (i.e., the learning rate). In both implementations, salience within each event across the two choice sets would change only marginally, where event $\mathcal{E}_3$ would receive the highest decision weight, failing to predict the choice pattern observed in choice set $\mathcal{S}_1$.

Within the reinforcement-learning framework, a different approach to context-dependent preferences is contextual value adaptation (Palminteri et al., 2015). In the present experimental design, contextual value would not exert any influence as our design explicitly controlled for contextual value, where both decoy options had the same expected value. Non-reinforcement-learning models often rely on recall of instances from memory to form preferences (Erev & Roth, 2014; Gonzalez & Dutt, 2011). Within these frameworks, individuals draw a sample of single trials from memory, process that sample, and choose the option with the highest criterion. These models could be augmented with a mechanism resembling the similarity mechanism in various different ways. For example, the values could be processed during the trial and these processed values (that already take into account outcome salience) could be stored in memory, people could recall an entire trial and then process it much like the reinforcement-learning model does. Irrespective of the concrete mechanistic implementation, the main phenomenon remains: Outcome salience in a context with relevant-but-invalid information affects preferences.

References

Berkowitsch, N. A. J., Scheibehenne, B., & Rieskamp, J. (2014). Rigorously testing multi-alternative decision field theory against random utility models. *Journal of Experimental Psychology: General*, *143*(3), 1331–1348, https://doi.org/10.1037/a0035159.

Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica*, *22*(1), 23–36, https://doi.org/10.2307/1909829.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2012). Salience theory of choice under risk. *The Quarterly Journal of Economics*, *127*(3), 1243–1285, https://doi.org/10.1093/qje/qjs018.

Busemeyer, J. R., Gluth, S., Rieskamp, J., & Turner, B. M. (2019). Cognitive and neural bases of multi-attribute, multi-alternative, value-based decisions. *Trends in Cognitive Sciences*, *23*(3), 251–263, https://doi.org/10.1016/j.tics.2018.12.003.

https://doi.org/10.1017/S1930297500009177 Published online by Cambridge University Press

Cavanagh, J. F., Wiecki, T. V., Kochar, A., & Frank, M. J. (2014). Eye tracking and pupillometry are indicators of dissociable latent decision processes. *Journal of Experimental Psychology: General*, *143*(4), 1476–1488, https://doi.org/10.1037/a0035813.

Erev, I. & Roth, A. E. (2014). Maximization, learning, and economic behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(Supplement_3), 10818–10825, https://doi.org/10.1073/pnas.1402846111.

Ert, E. & Lejarraga, T. (2018). The effect of experience on context-dependent decisions. *Journal of Behavioral Decision Making*, *31*(4), 535–546, https://doi.org/10.1002/bdm.2064.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. CRC Press, 3rd ed. edition.

Gluth, S., Spektor, M. S., & Rieskamp, J. (2018). Value-based attentional capture affects multi-alternative decision making. *eLife*, *7*, 1–36, https://doi.org/10.7554/eLife.39659.

Gonzalez, C. & Dutt, V. (2011). Instance-based learning: Integrating sampling and repeated decisions from experience. *Psychological Review*, *118*(4), 523–551, https://doi.org/10.1037/a0024558.

Hadar, L., Danziger, S., & Hertwig, R. (2018). The attraction effect in experience-based decisions. *Journal of Behavioral Decision Making*, *31*(3), 461–468, https://doi.org/10.1002/bdm.2058.

Herne, K. (1999). The effects of decoy gambles on individual choice. *Experimental Economics*, *2*, 31–40, https://doi.org/10.1023/A:1009925731240.

Hertwig, R. & Erev, I. (2009). The description–experience gap in risky choice. *Trends in Cognitive Sciences*, *13*(12), 517–523, https://doi.org/10.1016/j.tics.2009.09.004.

Huber, J., Payne, J. W., & Puto, C. P. (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.

Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–292, https://doi.org/10.2307/1914185.

Kellen, D. (2019). A model hierarchy for psychological science. *Computational Brain & Behavior*, *2*(3-4), 160–165, https://doi.org/10.1007/s42113-019-00037-y.

Kool, W., Cushman, F. A., & Gershman, S. J. (2016). When does model-based control pay off? *PLOS Computational Biology*, *12*(8), e1005090, https://doi.org/10.1371/journal.pcbi.1005090.

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. Wiley.

Müller-Trede, J., Sher, S., & McKenzie, C. R. M. (2015). Transitivity in context: A rational analysis of intransitive choice and context-sensitive preference. *Decision*, *2*(4), 280–305, https://doi.org/10.1037/dec0000037.

Navarro, D. J., Newell, B. R., & Schulze, C. (2016). Learning and choosing in an uncertain world: An investigation of the explore–exploit dilemma in static and dynamic environments. *Cognitive Psychology*, *85*, 43–77, https://doi.org/10.1016/j.cogpsych.2016.01.

https://doi.org/10.1017/S1930297500009177 Published online by Cambridge University Press

001.

Noguchi, T. & Stewart, N. (2018). Multialternative decision by sampling: A model of decision making constrained by process data. *Psychological Review*, *125*(4), 512–544, https://doi.org/10.1037/rev0000102.

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

Plonsky, O. & Erev, I. (2017). Learning in settings with partial feedback and the wavy recency effect of rare events. *Cognitive Psychology*, *93*, 18–43, https://doi.org/10.1016/j.cogpsych.2017.01.002.

Rakow, T., Newell, B. R., & Wright, L. (2015). Forgone but not forgotten: The effects of partial and full feedback in "harsh" and "kind" environments. *Psychonomic Bulletin & Review*, *22*(6), 1807–1813, https://doi.org/10.3758/s13423-015-0848-x.

Rieskamp, J., Busemeyer, J. R., & Mellers, B. A. (2006). Extending the bounds of rationality: Evidence and theories of preferential choice. *Journal of Economic Literature*, *44*(3), 631–661, https://doi.org/10.1257/jel.44.3.631.

Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionist model of decision making. *Psychological Review*, *108*(2), 370–392, https://doi.org/10.1037/0033-295X.108.2.370.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599, https://doi.org/10.1126/science.275.5306.1593.

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.

Soltani, A., De Martino, B., & Camerer, C. (2012). A range-normalization model of context-dependent choice: A new model and evidence. *PLOS Computational Biology*, *8*(7), 1–15, https://doi.org/10.1371/journal.pcbi.1002607.

Spektor, M. S., Bhatia, S., & Gluth, S. (2021). The elusiveness of context effects in decision making. *Trends in Cognitive Sciences*, *25*(10), 844–857, https://doi.org/10.1016/j.tics.2021.07.011.

Spektor, M. S., Gluth, S., Fontanesi, L., & Rieskamp, J. (2019). How similarity between choice options affects decisions from experience: The accentuation-of-differences model. *Psychological Review*, *126*, 52–88, https://doi.org/10.1037/rev0000122.

Sutton, R. S. & Barto, A. G. (1998). *Reinforcement learning: An introduction*. MIT Press.

Szollosi, A. & Newell, B. R. (2020). People as intuitive scientists: Reconsidering statistical explanations of decision making. *Trends in Cognitive Sciences*, *24*(12), 1008–1018, https://doi.org/10.1016/j.tics.2020.09.005.

Trueblood, J. S., Brown, S. D., Heathcote, A., & Busemeyer, J. R. (2013). Not just for consumers: Context effects are fundamental to decision making. *Psychological Science*, *24*(6), 901–908, https://doi.org/10.1177/0956797612464241.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*,

https://doi.org/10.1017/S1930297500009177 Published online by Cambridge University Press

79(4), 281–299, https://doi.org/10.1037/h0032955.

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131, https://doi.org/10.1126/science.185.4157.1124.

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413–1432, https://doi.org/10.1007/s11222-016-9696-4.

von Neumann, J. & Morgenstern, O. (1947). *Theory of games and economic behavior*. MIT Press, 2nd ed. edition.

Wedell, D. H. (1991). Distinguishing among models of contextually induced preference reversals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(4), 767–778, https://doi.org/10.1037//0278-7393.17.4.767.

Wulff, D. U., Mergenthaler-Canseco, M., & Hertwig, R. (2018). A meta-analytic review of two modes of learning and the description-experience gap. *Psychological Bulletin*, *144*(2), 140–176, https://doi.org/10.1037/bul0000115.

Yechiam, E. & Busemeyer, J. R. (2005). Comparison of basic assumptions embedded in learning models for experience-based decision making. *Psychonomic Bulletin & Review*, *12*(3), 387–402, https://doi.org/10.3758/BF03193783.

Yechiam, E. & Rakow, T. (2012). The effect of foregone outcomes on choices from experience. *Experimental Psychology*, *59*(2), 55–67, https://doi.org/10.1027/1618-3169/a000126.

https://doi.org/10.1017/S1930297500009177 Published online by Cambridge University Press