Author/s:
Voigt, K;Murawski, C;Bode, S

Title:
Endogenous formation of preferences: Choices systematically change willingness-to-pay for goods

Date:
2017-12

Citation:
Voigt, K., Murawski, C. & Bode, S. (2017). Endogenous formation of preferences: Choices systematically change willingness-to-pay for goods. Journal of Experimental Psychology: Learning, Memory, and Cognition, 43 (12), pp.1872-1882. https://doi.org/10.1037/xlm0000415.

Persistent Link:
https://hdl.handle.net/11343/300041

Journal of Experimental Psychology: Learning, Memory, and Cognition

Endogenous Formation of Preferences: Choices Systematically Change Willingness-to-Pay for Goods

Katharina Voigt, Carsten Murawski, and Stefan Bode

Online First Publication, May 15, 2017. http://dx.doi.org/10.1037/xlm0000415

Journal of Experimental Psychology:
Learning, Memory, and Cognition

0278-7393/17/$12.00 http://dx.doi.org/10.1037/xlm0000415

Endogenous Formation of Preferences: Choices Systematically Change Willingness-to-Pay for Goods

Katharina Voigt, Carsten Murawski, and Stefan Bode
The University of Melbourne

Standard decision theory assumes that choices result from stable preferences. This position has been challenged by claims that the act of choosing between goods may alter preferences. To test this claim, we investigated in three experiments whether choices between equally valued snack food items can systematically shape preferences. We directly assessed changes in participants' willingness-to-pay for these items, some of which could be bought at an auction after the experiment, while others could not. We found that chosen items were valued higher, and nonchosen items were valued lower; yet this postdecisional refinement of preferences was only observed for choices and valuations that were relevant, that is, incentive-compatible for items that were available for consumption. Supplementary analyses revealed that incentive-incompatible elicitations of preferences were unreliable and may have masked potential effects of choices on preferences. In conclusion, we propose that preferences can change endogenously, that is, in the absence of external feedback or information, but rather as a function of previous relevant choices.

Keywords: decision-making, choice-induced preference change, preference formation, valuation, willingness-to-pay

Supplemental materials: http://dx.doi.org/10.1037/xlm0000415.supp

Do we choose what we prefer, or do we prefer what we choose? At first glance, the answer to this question seems to be simple: standard decision theory assumes that our choices reflect stable preferences over available choice options (Samuelson, 1938). An alteration of preferences is only expected if new (external) information about choice alternatives becomes available (e.g., through the consumption of a good). However, it has been argued that the relationship between preferences and observed behavior might not be unidirectional (Bettman, Luce, & Payne, 1998; Fischhoff, 1991; Kahneman, Slovic, & Tversky, 1982; Thaler, 1980). Preferences may instead be context-sensitive and depend on how the choice problem is presented (Slovic, 1995; Tversky & Kahneman, 1986) and processed by the decision maker (Kahneman, 2003; Simon, 1955). Others have suggested that preferences can change endogenously even in the absence of any new information as a consequence of the mere act of choosing an option. For example, Brehm (1956) used the so-called "free-choice paradigm" in which participants first indicated their preference for a series of items; they then made binary decisions between choice pairs of equally preferred items; finally, they reported their preference for the same items again (valuation-choice-valuation sequence; V-C-V). Brehm (1956) showed that from the initially equally valued items, the chosen items became more valuable and the nonchosen items less valuable. A number of studies replicated this finding adopting the free-choice paradigm (e.g., Kitayama, Chua, Tompson, & Han, 2013; Lieberman, Ochsner, Gilbert, & Schacter, 2001; Sharot, De Martino, & Dolan, 2009; reviewed by Izuma & Murayama, 2013).

However, recent work has cast doubt on the validity of these findings. Chen and Risen (2010) argued that if valuations of items—as well as choices between items—are only noisy representations of underlying preferences, then each valuation (and choice) can be conceived as an independent draw from the same distribution. The observed spread of alternatives in the second valuation could simply reflect a naturally occurring regression-to-the-mean effect (that is, a regression to the true preference). Both valuations and the (value-based) choice together would then provide a more accurate reflection of the initial preferences over items that, in reality, were already different, rather than reflecting a choice-induced *change* of preferences between the two subsequent valuations (Chen & Risen, 2010). In consequence, these authors suggested employing a "valuation-valuation-choice" (V-V-C) con-

Katharina Voigt, Melbourne School of Psychological Sciences, The University of Melbourne; Carsten Murawski, Department of Finance, The University of Melbourne; Stefan Bode, Melbourne School of Psychological Sciences, The University of Melbourne.

This work was supported by the Strategic Initiatives Fund of the Faculty of Business and Economics, The University of Melbourne, to Carsten Murawski and Stefan Bode, and an Australian Research Council Discovery Early Career Researcher Award (DE 140100350) to Stefan Bode. We thank Piers Howe, Carmen Morawetz, Jutta Stahl, and Simon Lilburn for feedback and comments on the manuscript, and Nicole Stefanac, William Turner, Megan Edelman, and Hayley Warren for help with data acquisition. Katharina Voigt, Carsten Murawski, and Stefan Bode designed experiments; Katharina Voigt ran experiments and analyzed data; Katharina Voigt, Stefan Bode, and Carsten Murawski wrote the manuscript. All authors gave final approval for publication.

Correspondence concerning this article should be addressed to Katharina Voigt, Melbourne School of Psychological Sciences, The University of Melbourne, Parkville, VIC 3010, Australia. E-mail: kvoigt@pgrad.unimelb.edu.au

trol sequence, which allows for subtracting out valuation changes that were not related to the nested choice. Studies using this design have found either no evidence for choice-induced preference changes (Chen & Risen, 2010; Coppin et al., 2014—Experiment 1), or only effects for either rejected items for which valuations decreased (Coppin et al., 2014—Experiment 2; Izuma et al., 2010; Salti, Karoui, Maillet, & Naccache, 2014) or only for chosen items for which valuations increased (Koster, Duzel, & Dolan, 2015; Sharot, Fleming, Yu, Koster, & Dolan, 2012). Hence, whether and how preferences are shaped by preceding choices still remains an open question.

A second problem is that previous studies have exclusively assessed the value of items using *liking* ratings, or rankings (e.g., Chen & Risen, 2010; Coppin et al., 2014; Izuma et al., 2010; Koster et al., 2015; Salti et al., 2014; Sharot et al., 2012). This form of preference assessment comes with two interrelated problems: First, liking ratings might be susceptible to measurement noise (Murphy, Allen, Stevens, & Weatherhead, 2005), which can also induce regression-to-the-mean effects (Izuma & Murayama, 2013). More importantly, liking ratings are difficult to interpret. In real life, valuations are usually reflected in purchase commitments for

actually available goods. These commitments, in turn, make the valuations consequential and incentive-compatible. Indeed, the concept of preferences has mostly been studied using nonincentive compatible preference assessments in form of the above-mentioned liking rating scales. As these statements assess valuations but do not have any direct consequences for obtaining the evaluated outcome, they also do not allow for reliably quantifying valuations with respect to a personal investment to obtain a good. Liking ratings might significantly deviate from such incentive-compatible valuations (Cummings, Harrison, & Rutstrom, 1995; Johannesson, Lilijas, & Johansson, 1998; Murphy et al., 2005). Note that even in the few studies that have implemented the revised free-choice paradigm, participants' preferences were still assessed using (hypothetical) liking ratings (e.g., Coppin et al., 2014; Sharot et al., 2012). In some studies participants could indeed receive an item they liked (e.g., Chen & Risen, 2010; Izuma et al., 2010; Koster et al., 2015), which has been suggested to increase task-alertness and truthfulness of responses for popular items (Murphy et al., 2005). However, even this scenario does not establish incentive-compatible valuations for *all* presented items, as the liking ratings are still not statements of preferences that reflect real—and potentially consequential—commitments.

In contrast, incentive-compatible measures of preferences directly relate the valuation itself to an outcome. One example for this is assessing the willingness-to-pay for items in combination with auction procedures for which limited monetary resources are available. In consequence, incentive-compatible valuation measures are typically more precise and immediate, have higher stakes, and are often more emotionally charged, whereas incentive-incompatible choices have no consequence, might be rapid, "mindless" and require fewer cognitive resources (Camerer & Mobbs, 2017; Kang & Camerer, 2013; Morgenstern, Heldmann, & Vogt, 2014; Morkbak, Olsen, & Campbell, 2014).

In this study, we assessed choice-induced preference changes via participants' stated willingness-to-pay (WTP) for consumable goods (snack foods) in an incentive-compatible fashion. This allowed us to directly compare the effects to (hypothetical) liking assessments of preferences. We further manipulated the incentive for making real decisions in the WTP condition by making half of the items potentially available for purchase and consumption after the experiment. The WTP condition was followed by an auction (Becker, DeGroot, & Marschak, 1964) in which participants' bids for the available and chosen items determined whether they received one item and could consume it. In order to eliminate regressions to the true preference (henceforth referred to as regression-to-the-mean effect), a V-V-C control sequence was implemented (Chen & Risen, 2010). In a second experiment, two independent samples stated their WTP for either only available or only unavailable items to investigate the robustness of choice-induced preference change effects for differently incentivized choices.

Method

Experiment 1

Participants. The first experiment was split into two experimental conditions with independent samples: 40 participants were randomly assigned to the *liking rating* condition and 40 participants were assigned to the *willingness-to-pay* (WTP) condition. After excluding one participant, who indicated not understanding the task procedures (see supplementary material, "Control questions assessing task understanding in the willingness-to-pay condition") and five participants after technical problems with data recording, 39 participants were included in the *liking rating* condition (24 females, mean age = 21.48, *SD* = 2.90 years) and 35 participants in the WTP condition (24 females, mean age = 22.31, *SD* = 2.92 years). Sample size was chosen based on a previous study (Izuma et al., 2010), which used a similar revised free-choice paradigm and stimuli as reported here. All participants were right-handed and had normal or corrected-to-normal visual acuity and no known history of neurological and psychiatric disorders. The two groups did not differ in any demographic and cognitive variables assessed (Table S1). In order to increase the salience of the experimental stimuli, all participants were asked to refrain from eating or drinking (except for water) for four hours prior to the study (Schonberg et al., 2014). Participants were recruited via advertisements at the University of Melbourne, gave written informed consent to the study, and were reimbursed with AUD 20. The study was approved by the University of Melbourne Human Research Ethics Committee (no. 1442440) and was conducted according to the Declaration of Helsinki.

Experimental stimuli. Stimuli consisted of 240 images of savory and sweet snack foods that were available in local grocery stores in Melbourne, Australia. Stimuli were selected via an online pilot study (see supplementary online material, "Online study for the selection of stimuli") completed by an independent group of 297 individuals (207 females, mean age = 26.01, *SD* = 7.74 years). Criteria were positive liking ratings, high familiarity, stated tastiness, on average purchased at least once per month, and an average stated WTP that did not exceed AUD 5. From all snack food stimuli meeting these criteria, 80 images were randomly selected for the Liking condition, and 160 items were selected for the WTP condition. In the latter condition, items were randomly distributed between the WTP-na (na—not available) condition (80 items), in which items were not available for consumption, and the

WTP-a (a—available) condition (80 items), in which items were available for consumption. Stimulus characteristics did not differ across conditions (Table S2). All stimuli were sized maximally 500 pixels in width or height and were presented on a gray background on a computer screen. Participants were seated in a comfortable chair, at a distance of 110 cm from the monitor (visual angle of 13.10° × 8.67°). Stimulus presentation and data acquisition were performed using MATLAB (R2012b; MathWorks) in combination with the Psychophysics toolbox (www.psychtoolbox.org).

Experimental paradigm. Prior to the experiment, participants' hunger state was assessed on an 8-point rating scale (1 = not hungry at all; 8 = very hungry), and it was confirmed that participants had fasted for at least 4 hours (see Table S1 for summary statistics). Participants in the WTP condition were further informed that only some snack food items would be available for an auction after the experiment (i.e., items shown in the WTP-a condition), and that one of these available items from the pool of items chosen during the experiment would randomly be selected for the auction. Participants received AUD 4 of bidding money. They were told to use this money to indicate their WTP for each of the presented items, and that the specific amounts would determine the probability of obtaining the item via a Becker-DeGroot-Marschak auction (BDM; for details see below; Becker et al., 1964). The experiment comprised four task phases (Figure 1).

(i) First valuation phase. Each of the 80 trials in the Liking condition started with a central fixation cross (jittered duration of 1.5 s to 5.5 s), which was followed by a pseudorandomly selected item (duration 1 s). Subsequently, participants indicated their liking for that item on a continuum from 1 (do not like it at all) to 8 (like it very much). All 80 items were shown once during this task phase. The WTP condition comprised 160 trials showing stimuli that were either available (80 WTP-a trials) or not available (80 WTP-na trials) for consumption after the experiment. The WTP-a and WTP-na trials were presented in a pseudorandom fashion. Each trial started with a central fixation cross (jittered duration of 1 s to 4 s), followed by a short text indicating whether the following item would not be available (in WTP-na trials) or would be available (in WTP-a trials; 1.5 s). This was followed by the presentation of one pseudorandomly selected item (stimulus duration 1 s) until all 160 items (80 in WTP-na, 80 in WTP-a) were shown once. For each item, participants indicated how much between AUD 0 to AUD 4 they would *hypothetically* be willing to spend on (bid for) an item (WTP-na), and how much they would *actually* be willing to spend on an item (WTP-a) in the following auction. Responses were measured by moving a graphical slider along a continuous valuation scale. The slider was controlled by a track pad, which was operated using the right hand. Previous studies demonstrated that continuous graphical sliders reduce measurement error compared to Likert scales (Treiblmaier & Filzmoser, 2011). There was no response time restriction.

(ii) First two-alternative forced-choice (2AFC) phase. In all conditions, each trial started with a short fixation period (jittered between 1.5 s and 4 s, followed by 1.5 s fixation cross in the Liking condition, or item availability information in the WTP condition). In each trial, participants in the liking and WTP condition were then presented with an item pair (one item presented on each side of the central fixation cross), generated from the items shown in the first valuation phase, and they were asked to choose the item they preferred out of these two items. They indicated their choice using a left and a right button on the keyboard, which terminated the trial. A maximum of 22 "difficult" and 13 "easy" choice pairs were created for the Liking condition and each trial type of the WTP condition (WTP-na, WTP-a) based on the responses in the

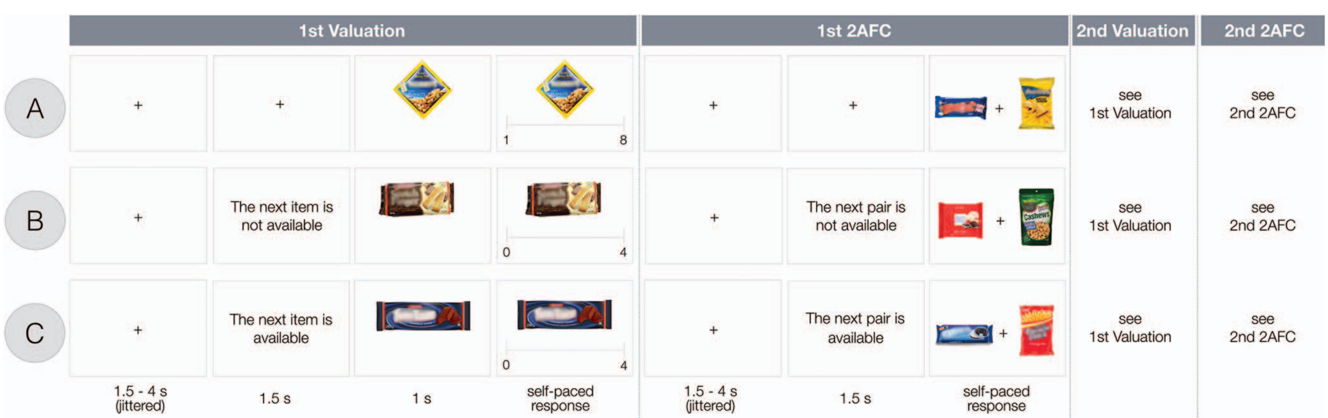

Figure 1. Paradigm main experiment. All experimental conditions comprised four task phases: First valuation, first two-alternative forced-choice (2AFC), second valuation, and second 2AFC. The second 2AFC phase with different choice pairs allowed for the implementation of a control sequence (valuation-valuation-choice), which controlled for regression-to-the-mean (see text). (A) Trial sequence in Liking rating condition. Liking ratings were given on a scale from 1 (do not like it at all) to 8 (like it very much). Items were never available for purchase; participants received an item they liked at the end of the experiment. (B) Trial sequence in Willingness-to-pay (WTP) condition with nonavailable items. Valuation was indicated as the hypothetical WTP for an item between AUD 0 and AUD 4. Items were never available for purchase. (C) Trial sequence in WTP condition with available items. Valuation was indicated as the actual WTP for an item between AUD 0 and AUD 4. Items were available for purchase, and based on decisions and valuations in the experiment, one item was later selected for an auction. See the online article for the color version of this figure.

first valuation phase, by pairing items with either highly similar (difficult) valuations or dissimilar (easy) valuations, respectively. We included a higher number of difficult choice pairs compared to the number of easy choice trials, on the basis of previous studies (e.g., Coppin et al., 2014; Izuma et al., 2010; Koster et al., 2015; Sharot, Velasquez, & Dolan, 2010). It is well-established that choice-induced preference effects only occur when individuals have to choose between equally valued options (reviewed by Izuma & Murayama, 2013). Hence, in order to achieve a higher statistical power as well as comparability with previous studies, the ratio of difficult to easy choice pairs was kept at 70/30. Half of these difficult and easy choice pairs were presented in this first 2AFC phase in a pseudorandomized order. Notably, for the construction of the difficult choice pairs, only those items were selected that were (a) valued above the individual's median valuation in order to render them desirable, and (b) had a minimum valuation difference between them. If choice pairs resulted for which the valuation difference exceeded 1 *SD* of the participant's average valuation difference, this pair was considered too dissimilar and excluded from all analyses. For easy choice pairs, items from the remaining item pool were paired such that the valuation differences were maximized. An average number of 18 choice pairs were created for the Liking condition in Experiment 1. An average number of 18.43 choice pairs were created within the WTP-na condition, and an average of 18.77 choice pairs were randomly created within the WTP-a condition (Table S3). These choice pairs were then randomly allocated to either the first or second 2AFC task stage within each experimental condition. Table S3 provides an overview of the choice trials that were included in the reported analyses.

(iii) Second valuation phase. This task phase was identical to the first valuation phase. Participants were instructed that the purpose was not to probe their memory of the first valuation, but to provide another, independent valuation (adopted from Sharot et al., 2009).

(iv) Second 2AFC phase. This task phase was identical to the first 2AFC phase with the only difference being that the remaining halves of the 20 difficult and easy choice pairs were presented that were not used in the first 2AFC phase. This second 2AFC phase with unused choice pairs allowed for using the first and second valuations for these items as a control sequence (V-V-C), assessing changes in valuation that could not be driven by choice, but were attributable to regression-to-the-mean effects (Chen & Risen, 2010).

After the experiment, a BDM auction (Becker et al., 1964) was conducted for the participants in the WTP condition. Following the procedure described by Plassmann, O'Doherty, and Rangel (2007), for each of the available and chosen items in task phase the bidding value *b* was compared to a value *n,* randomly drawn from a uniform distribution of all possible bidding values (AUD 0 to AUD 4, in increments of 5 cents). If *b* exceeded or was equal to *n,* the bid was considered successful, and the participant received the item for a price equal to *n*. If *b* was smaller than *n,* the participant did not receive the item and kept the preallocated money instead. This procedure was known to the participants and ensured that bids and decisions for available items during the experiment were incentivized and consequential. Only participants who completely understood this procedure (assessed via control questions; see supplementary online material, "Control questions assessing task understanding in the willingness-to-pay condition") were considered for experimental testing. Note that this implementation of a BDM auction makes the WTP condition incentive-compatible, as the optimal strategy in a BDM auction is to bid the true valuation for each item. The reason for this is that there is no incentive to bid less than the WTP, because the final price to be paid is determined by a random number *n* and does not affect the bid. There is no incentive to bid above the WTP either, because in this situation the individual might win the item but would end up paying an inflated price. The fact that bidding the WTP is the optimal strategy was explained and emphasized during the instruction period. Finally, participants filled out questionnaires regarding their diet.

Data analysis. For each participant and each item, the change in valuation (liking rating or WTP) was calculated by subtracting the first valuation from the second valuation. In order to correct for global shifts in valuation (e.g., participants might be more likely to indicate higher (or lower) valuations overall in the first valuation compared to the second valuation), only mean-corrected scores were used by first subtracting the average bid for the respective individual participant's session from the raw score of a single trial ($xi - mean_{session}$; Salti et al., 2014; Sharot et al., 2009). Following Sharot and colleagues (2009), the $mean_{session}$ included all the valuations for all items, regardless of the later allocated condition (i.e., chosen-difficult, chosen-easy, rejected-difficult, rejected-easy). For the liking rating condition, these difference valuation scores were then subjected to a repeated-measures analysis of variance (ANOVA) with the within-subject factors Choice Outcome (chosen, rejected) and Experimental Sequence (V-C-V, V-V-C). For the WTP condition, a similar repeated-measures ANOVA was conducted with the within-subject factors Choice Outcome (chosen, rejected), Experimental Sequence (V-C-V, V-V-C), and Item Availability (not available, available). True choice-induced preference changes, corrected for regression-to-the-mean effects, required *differences* in valuation changes between experimental sequences to be significant. As choice-induced preference change effects can only be expected for difficult trials (Chen & Risen, 2010; Coppin et al., 2014; Izuma et al., 2010; Salti et al., 2014), the key analyses were conducted for the difficult trials.

Experiment 2

Potential choice-induced preference change effects for nonavailable items in the WTP-na condition (as observed in our study; see Results) could potentially be explained by the fact that in our experiment available and nonavailable items were presented in a randomized order. That is, participants might apply the same task strategy in valuating available and nonavailable items independently. In order to investigate whether nonavailable items elicited a choice-induced preference change effect, even when participants were not presented with available items during the experiment, a second experiment was conducted. This experiment replicated the WTP condition of the main experiment, but the WTP-a trials and WTP-na trials were now shown in a between-subjects design.

Participants. Two different, independent samples of participants (Table S4) were assigned to each condition and presented with the same task in which *either* only nonavailable items (WTP-na) *or* only available items (WTP-a) were used. The sample size was 33 participants for the WTP-na condition (25 female, mean age = 22.97, *SD* = 4.25 years), and 33 participants for the WTP-a condition (19 females, mean age = 22.28, *SD* = 3.43 years). All participants were right-handed, had normal or corrected-to-normal visual acuity, and no known history of neurological and psychiatric disorders. Participants were again recruited via advertisements at the University of Melbourne, gave written informed consent to the study, and were reimbursed with AUD 20.

Experimental stimuli and paradigm. One hundred and sixty stimuli were randomly drawn from the same item pool that was used for the first experiment, and randomly allocated to the WTP-na (80 items) and the WTP-a (80 items) conditions, such that they did not differ on any attribute across conditions (Table S5). The experimental procedure for both conditions was identical to the WTP conditions of the first experiment. A BDM auction (Becker et al., 1964) was conducted, as described for the first experiment, for the WTP-a group only. All other procedures were identical to the first experiment.

Data analysis. For each of the two conductions, a two-way repeated-measures ANOVA with Choice Outcome (chosen, rejected) and Experimental Sequence (V-C-V, V-V-C) as within-subject factors was conducted to analyze the data.

Experiment 3

A final experiment was conducted to replicate the results of the Liking condition in an independent sample of 35 participants (20 females, mean age = 27.60, *SD* = 6.16 years). This experiment further controlled for the possibility that the use of different scale labels (Experiment 1 used 1 to 8 for Liking ratings, but 0 to 4 for WTP) might have biased the results. Hence, this experiment used 0 to 4 as endpoint labels of the response scale. All other aspects of the experiment were identical to the Liking condition of Experiment 1 (see supplementary material, "Liking Control Experiment," for details).

Results

Experiment 1

Valuations and choices. In the Liking condition, participants rated the items on average at 5.49 (*SD* = 2.10, range 1–8), which was slightly above the average of the liking scale. Note that items were selected in pretests such that they were likable, and participants were familiar with them (Table S2). The mean liking rating was significantly lower in the second valuation phase ($M = 5.41$, $SD = .65$) compared to the first ($M = 5.50$, $SD = .62$), $t(38) = 2.31$, $p = .03$; Table S7. However, as all values were mean-corrected for each valuation stage for our main analyses (Sharot et al., 2009), this difference could not affect the choice-induced-preference change analysis. Across all items and all participants, the correlation between liking ratings from the first and second valuation stage was low but significant (Pearson's $r = .13$, $p = .01$, 95% CI_{boot} = [.02, .24]; Figure S1). In the 2AFC tasks, for difficult choice pairs, the choice ratio for the initially higher valued item (mean ratio = .39, *SD* = .19) was significantly lower than for the initially lower valued items (mean ratio = .65, *SD* = .15; $t(38) = -7.5$, $p < .001$; Table S8).

In the WTP condition, participants' average bidding amount was AUD 1.71 (*SD* = .53, range 0–4) for unavailable items and AUD 1.78 (*SD* = .52, range 0–4) for available items, covering the entire bidding range (Table S7). The mean bids did not differ between the first and second valuation stages within conditions (WTP-na: $t(34) = .17$, $p = .87$; WTP-a: $t(34) = .14$, $p = .89$), and also did not differ between the WTP-na and the WTP-a conditions (first valuation phase: $t(34) = 1.8$, $p = .08$; second valuation phase: $t(34) = 1.7$, $p = .10$). This means that it was unlikely that different bidding behavior occurred in the WTP conditions, although the *p*-value for the first valuation stage was rather small. Correlations between first and second bids were significant for both unavailable items (WTP-na: Pearson's $r = .80$, $p < .001$, 95% CI_{boot} = [.76, .83]) and available items (WTP-a: Pearson's $r = .87$, $p < .001$, 95% CI_{boot} = [.02, .24]; Figure S1). For difficult choice pairs in the 2AFC, items for which a slightly higher WTP was indicated in the first valuation were not chosen more often than items for which a slightly lower WTP was indicated, neither in the WTP-na condition, $t(34) = -.10$, $p = .92$, nor in the WTP-a condition, $t(34) = -.10$, $p = .19$, indicating a successful pairing of equally valued choice pairs (Table S8).

Choice-induced preference change. First, we established whether the second valuation changed from the first valuation for each experimental condition (Liking, WTP-na, WTP-a) in the classical valuation-choice-valuation (V-C-V) sequence. Using a repeated-measures ANOVA for the mean valuation change scores with Choice Outcome (chosen, rejected) as a within-subjects factor, a significant main effect was found ($F(1, 38) = 57.25$, $p < .001$, $\eta^2 = .60$). The liking ratings for rejected items significantly decreased, $t(38) = -9.42$, $p < .001$, Cohen's $d = 1.51$. This alone might only be indicative for a regression-to-the-mean effect. Very similar rejection effects were found in previous studies (e.g., Coppin et al., 2014; Izuma et al., 2010; Salti et al., 2014). There was no significant change observed for chosen items $t(38) = -1.13$, $p = .26$ (see Figure 2). For the WTP conditions, a repeated-measures ANOVA with Choice Outcome (chosen, rejected) and Item Availability (not available: WTP-na, available: WTP-a) as within-subject factors showed significant main effects for Choice Outcome ($F(1, 34) = 178.28$, $p < .001$, $\eta^2 = .84$) but not for Item Availability, $F(1, 34) = .28$, $p < .60$, and no significant interaction, $F(1, 34) = .23$, $p < .64$. Post hoc *t*-tests showed a significant increase in bids in the postchoice valuation phase for chosen items in both the WTP-na and the WTP-a conditions (WTP-na: $t(34) = 3.99$, $p < .001$, Cohen's $d = .67$; WTP-a: $t(34) = 5.29$, $p < .001$, Cohen's $d = .89$), as well as a significant decrease in bids for rejected items (WTP-na $t(34) = -9.70$, $p < .001$, Cohen's $d = 1.64$; WTP-a: $t(34) = -8.15$, $p < .001$, Cohen's $d = 1.38$). However, as argued by Chen and Risen (2010), these results alone are not informative about whether or not a true choice-induced preference change effect or a regression-to-the-mean effect was responsible for the observed valuation changes.

In order to investigate whether true choice-induced changes in preference/valuation occurred in each experimental condition, the difference between the first and second valuations were compared between the V-C-V sequence and the valuation-valuation-choice V-V-C sequence, which adjusted for potential regression-to-the-mean effects in the second valuation. This was performed separately for chosen items and rejected items. For the liking rating condition, using repeated-measures ANOVA with Experimental Sequence (V-C-V, V-V-C) and Choice Outcome (chosen, rejected) as within-subject factors, no main effect was found for experimen-

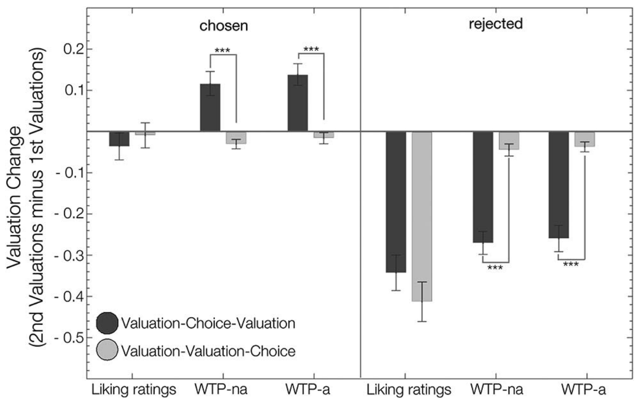

Figure 2. Results Experiment 1. Mean-corrected valuation changes for items, which were chosen (left panel) and rejected (right panel) in the respective 2AFC task phase. For each experimental condition (liking ratings, WTP-na, WTP-a), the valuation changes in the valuation-choice-valuation (V-C-V; dark gray) and the valuation-valuation-choice (V-V-C; light gray) sequences are displayed. Choice-induced preference change, corrected for regression-to-the-mean effects between valuations, is indicated as a significant difference in valuations changes between experimental sequences. Significant choice-induced preference change effects could be found for both WTP conditions, irrespectively of item availability. These effects were not found for the liking rating condition (see Table 1 for statistics). Note that because mean-corrected scores were calculated by subtracting the average bid across all trials from the respective session, and analyses were conducted on difficult trials only, the scores displayed do not average to zero. Valuations in the WTP conditions and the Liking condition were measured on different scales and cannot be compared directly. *** $p < .001$; error bars indicate *standard error of the mean (SEM)*.

tal sequence, $F(1, 38) = .01$, $p < .93$. The main effect for Choice Outcome was significant ($F(1, 38) = 104.90$, $p < .001$, $\eta^2 = .73$); however, the two-way interaction between Choice Outcome and Experimental Sequence was not significant, $F(1, 38) = 1.00$, $p = .32$, indicating that the observed change in ratings was not attributable to choice per se, but also occurred without the interspersed choice. This means that regression-to-the-mean can entirely explain the change, providing no evidence for the originally proposed choice-induced preference change effect for valuations via liking ratings.

For the WTP condition, a repeated-measures ANOVA was conducted, with Experimental Sequence (V-C-V, V-V-C), Choice Outcome (chosen, rejected), and Item Availability (WTP-na, WTP-a) as within-subject factors. A main effect of Choice Outcome ($F(1, 34) = 217.83$, $p < .001$, $\eta^2 = .87$) was found. There were no other significant main effects, but a significant two-way interaction effect for Choice Outcome and Experimental Sequence ($F(1, 34) = 10.61$, $p < .003$, $\eta^2 = .24$). Bonferroni-corrected repeated measures *t*-tests revealed that the increase in valuation change was significantly higher in the V-C-V sequence compared to the V-V-C sequence for chosen available items, $t(34) = 5.95$, $p < .001$, Cohen's $d = .76$, and also for chosen nonavailable items, $t(34) = 4.49$, $p < .001$, Cohen's $d = .78$. Similarly, there was a significant difference in decrease of valuations in the V-C-V sequence compared to the V-V-C sequence for both rejected available items, $t(34) = -7.13$, $p < .001$, Cohen's $d = 1$, and rejected nonavailable items, $t(34) = -7.40$, $p < .001$, Cohen's $d = .71$ (Table 1), indicating choice-induced preference change effects. Finally, there was no difference in the choice-induced preference change effect between the WTP-na and WTP-a conditions, as indicated by the absence of a significant three-way interaction, $F(1, 34) = .37$, $p = .55$ (see Figure 2). There were no other significant interaction effects.

We conducted additional repeated-measures ANOVAs for the easy choice trials. No interaction effects of choice outcome with experimental sequence were found for either condition (Liking condition: $F(1, 38) = .032$, $p < .86$; WTP condition: $F(1, 34) = .055$, $p < .82$). This result is consistent with previous reports suggesting that choice-induced preference change effects are restricted to cases of initial preference equality (Chen & Risen, 2010; Izuma et al., 2010).

Experiment 2

For Experiment 1, we initially hypothesized a stronger effect for incentivized valuations for available items (WTP-a) as compared to unavailable items (WTP-na). Our results did not support this hypothesis, as we found equally strong effects for both WTP conditions. However, the absence of this difference could be

Table 1
Comparisons of Valuation Changes for Chosen and Rejected Items in the Each Condition of Experiment 1

		V-C-V		V-V-C		Statistic		
Condition	Choice	Mean	*SD*	Mean	*SD*	*t*	*p*	Cohen's *d*
Liking ratings	Chosen	−.04	.23	−.01	.19	$t(34) = -.74$	.46[a]	
	Rejected	−.34	.27	−.41	.30	$t(34) = 1.31$	.40[a]	
WTP-na	Chosen	.12	.17	−.03	.07	$t(34) = 4.49$	<.001[b]	.78
	Rejected	−.27	.17	−.04	.07	$t(34) = -7.40$	<.001[b]	.71
WTP-a	Chosen	.14	.16	−.02	.08	$t(34) = 5.95$	<.001[b]	.76
	Rejected	−.26	.19	−.03	.07	$t(34) = -7.13$	<.001[b]	1

Note. V-C-V = valuation-choice-valuation sequence; V-V-C = valuation-valuation-choice sequence; *SD*, standard deviation; WTP-na = Willingness-to-pay (non-available items); WTP-a = Willingness-to-pay (available items). Statistic represents two-paired samples *t*-test between the mean valuation differential in the V-C-V sequence and the mean valuation differential in the V-V-C sequence.
[a] Bonferroni-corrected *p*-values for two tests. [b] Bonferroni-corrected *p*-values for four tests.

explained by our experimental design, as participants were presented with a randomized, intermixed sequence of all available and not available items. It is conceivable that participants treated all items the same, as if all items were available, regardless of the trial-by-trial indication of their true availability. The second experiment was designed to investigate with independent samples whether item availability moderated the choice-induced preference change effect, and participants were only exposed to *either* available *or* unavailable items.

Valuations and choices. Participants in the WTP-na condition indicated to be willing to pay AUD 1.91 on average (SD = 1.24, range 0–4), and participants in the WTP-a condition indicated to be willing to pay AUD 1.98 on average (SD = 1.13, range 0–4). The mean bids did not differ between the first and second valuation stages within conditions (WTP-na: $t(32) = 1.45$, p = .15; WTP-a: $t(32) = 1.84$, p = .88; Table S9). The mean bids between the WTP-na and the WTP-a conditions were also not significant in the first valuation phases, $t(32) = -.06$, p = .96, or second valuation stages of the two WTP conditions, $t(32) = -.42$, p = .68. Correlations between the first and second bids were significant for both unavailable items (WTP-na: Pearson's r = .53, $p < .001$, 95% CI_{boot} = [.50, .56]) and available items (WTP-a: Pearson's r = .70, $p < .001$, 95% CI_{boot} = [.68, .73]; Figure S2). For difficult choice pairs (Table S6), choice ratios for items with an initially slightly higher WTP were not significantly different from choice ratios for items with an initial slightly lower value, neither in the WTP-na condition, $t(32) = -.10$, p = .92, nor in the WTP-a condition, $t(32) = .14$, p = .89. This indicates a successful pairing of equally valued choice pairs (Table S10).

Choice-induced preference change. For each condition (WTP-na, WTP-a), repeated-measures ANOVAs were conducted, with Experimental Sequence (V-C-V, V-V-C) and Choice Outcome (chosen, rejected) as within-subject factors. For the WTP-na condition, there was a main effect of Choice Outcome (F (1, 32) = 62.38, $p < .001$, η^2 = .66). There was no significant main effect of experimental sequence, F (1, 32) = .08, $p < .79$. Although there was a significant interaction effect for Choice and Experimental Sequence (F(1, 32) = 4.65, $p < .04$, η^2 = .13), a Bonferroni-corrected paired sample t-test showed that the difference in valuation change between the V-C-V sequence and the V-V-C control sequence was not significant, neither for chosen items, $t(34)$ = 1.48, p = .15, nor for rejected items, $t(34) = -.39$, p = .70; Figure 3, Table 2). For the WTP-a condition, there was a significant main effect of Choice (F (1, 32) = 147.43, $p < .001$, η^2 = .82), no main effect of Experimental Sequence, F(1, 32) = .01, p = .69, but a significant interaction effect for Experimental Sequence and Choice (F(1, 32) = 20.08, $p < .001$, η^2 = .39). Bonferroni-corrected repeated measures t-tests showed that the valuation change for chosen items was significantly higher in the V-C-V sequence compared to the V-V-C sequence, $t(34)$ = 2.62, p = .03, Cohen's d = .46, and significantly lower for rejected items $t(34) = -2.41$, p = .03, Cohen's d = .42 (Figure 3, Table 2). In summary, a robust significant choice-induced preference effect was only observed for available items. As in Experiment 1, we found no evidence for a choice-induced preference change effect for the easy trials (i.e., no significant interaction effect for Experimental Sequence and Choice; WTP trials, F(1, 32) = .427, p = .56; WTP-na trials, F(1, 32) = .042, p = .83.

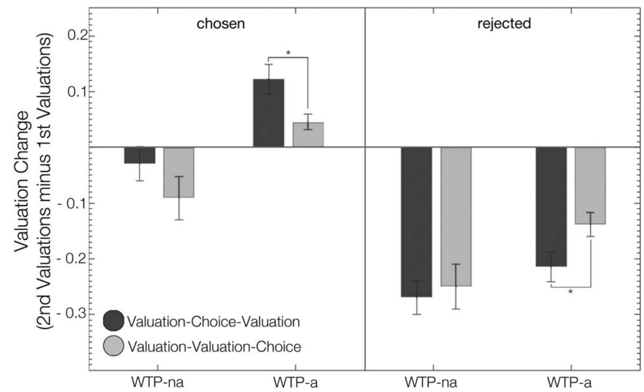

Figure 3. Results Experiment 2. Mean-corrected valuation changes for items, which were chosen (left) and rejected (right) in the respective 2AFC task phase. For each experimental condition (WTP-na, WTP-a), the valuation changes in the valuation-choice-valuation (V-C-V; dark gray) and the valuation-valuation-choice (V-V-C; light gray) sequences are displayed. Choice-induced preference change, corrected for regression-to-the-mean effects between valuations, is indicated as a significant difference in valuations changes between experimental sequences. Significant choice-induced preference change effects could be found for the WTP-a condition with available items. This effect was not present for the WTP-na condition with unavailable items (see Table 2 for statistics). Note that because mean-corrected scores were calculated by subtracting the average bid across all trials from the respective session, and analyses were conducted on difficult trials only, the scores displayed do not average to zero. * $p < .05$; error bars indicate SEM.

Experiment 3

Finally, a control experiment was conducted to replicate the results from the Liking condition in an independent sample, but used end-point labels of 0 to 4, identical to the WTP experiments. We observed the same pattern of results as in the Liking condition of Experiment 1 (for details regarding valuations and choice distributions see supplementary material: "Liking Control Experiment"). A repeated-measures ANOVA for the mean valuation change scores, with Choice Outcome (chosen, rejected) as the within-subjects factor, confirmed the significant main effect (F(1, 34) = 88.34, $p < .001$, η2 = .72), whereby values for rejected items significantly decreased, $t(48) = -10.23$, $p < .001$, but values for chosen items did not significantly increase, $t(34)$ = .23, p = .82 (Figure S4). In order to test for choice-induced preference change effects, we again used a repeated-measures ANOVA with experimental sequence (V-C-V, V-V-C) and choice outcome (chosen, rejected) as within-subject factors. As before, no main effect was found for experimental sequence, F(1, 34) = .90, p = .35. The main effect for choice outcome was significant (F(1, 34) = 88.34, $p < .001$, η2 = .72), but the two-way interaction between choice outcome and experimental sequence, which is indicative of choice-induced preference change effects, was again not significant, F(1, 34) = 1.82, p = .19. This indicated that the observed change in ratings was not attributable to choice per se, but again occurred without the interspersed choice (see Table S12 for additional post hoc comparisons). This means that regression-to-the-mean can entirely explain the change, as in the Liking condition of Experiment 1.

Table 2
Comparisons of Valuation Changes for Chosen and Rejected Items in the WTP-na and WTP-a Condition Experiment 2.

		V-C-V		V-V-C		Statistics		
Condition	Choice	Mean	*SD*	Mean	*SD*	*t*	*p*	Cohen's *d*
WTP-na	Chosen	−.03	.18	−.09	.22	$t(32) = 1.48$	.15[a]	
	Rejected	−.27	.17	−.25	.23	$t(32) = -.39$	.70[a]	
WTP-a	Chosen	.12	.15	.05	.09	$t(32) = 2.62$	.03[a]	.46
	Rejected	−.21	.15	−.14	.12	$t(32) = -2.41$	.03[a]	.42

Note. V-C-V = valuation-choice-valuation sequence; V-V-C = valuation-valuation-choice sequence; *SD*, standard deviation; WTP-na = Willingness-to-pay (non-available items); WTP-a = Willingness-to-pay (available items). Statistic represents two-paired samples *t*-test between the mean valuation differential in the V-C-V sequence and the mean valuation differential in the V-V-C sequence.
[a] Bonferroni-corrected *p*-values for two tests.

Discussion

In three experiments, we investigated the effect of previous choices on preferences in incentivized and non-incentivized decision scenarios. Our results show that valuations changed endogenously when valuations were incentive-compatible and assessed by willingness-to-pay, but not when assessed by hypothetical liking ratings. Effects in the liking rating condition were attributable to regression-to-the-mean effects (Chen & Risen, 2010). This was confirmed by a replication study using an independent sample that also modified the scale labels such that they were identical to the willingness-to-pay conditions. In the second experiment, we further demonstrated that choice-induced preference change effects were only robust when items were actually available for consumption later on, rendering choices and valuations relevant. These findings support the notion of choice-induced preference change effects, complementing previous studies, which used "blind choice" paradigms to demonstrate that choice itself can alter preferences (e.g., Egan, Bloom, & Santos, 2010; Sharot et al., 2010). We showed that if no clear preferences between relevant choice options preexisted, detecting equality between choice options may drive the preference system toward refinement, and the choice itself may be treated as a proxy for the anticipated consumption. In this model, preferences are dynamic, integrating behaviorally informed refinements and adjustments, allowing the organism to adapt to the environment and improve future outcomes.

Our study investigated whether preference formation can be endogenous, by using a variant of the so-called "free-choice paradigm" (Brehm, 1956). For this, we compared the spread of alternatives observed in the classical valuation-choice-valuation sequence (Brehm, 1956) to a control valuation-valuation-choice sequence. This made it possible to subtract out the effect of repeated valuation that was attributable to a regression to the mean (Chen & Risen, 2010). Previous studies, which implemented this control technique in combination with liking ratings, reported no effect (Chen & Risen, 2010; Coppin et al., 2014—their Study 1), effects only for chosen items (Koster et al., 2015; Sharot et al., 2012), or only for rejected items (Coppin et al., 2014—Study 2; Izuma et al., 2010; Salti et al., 2014). Our results found no choice-induced preference change effects when valuations were assessed using classical liking ratings.

Crucially, none of these previous studies used incentive compatible valuation assessments and choices with consequential commitments. The studies in which participants could receive an item that they liked at the end of the experiment (e.g., Chen & Risen, 2010; Izuma et al., 2010), however, did not use incentive-compatible valuations for all presented items. Our study addressed this problem by not only using item availability to make choices consequential, but by additionally implementing an incentive-compatible preference assessment, that is, the willingness-to-pay procedure. In the WTP condition, participants indicated their preferences similar to how everyday preferences are expressed—by allocating a fraction of an existing budget to the purchase of a good. Our supplementary analyses (available online) further showed that the correlations between the first and second valuation were lower for the liking ratings (also confirmed in our control study) compared to the WTP conditions, which could be a consequence of greater reliability of the WTP measures in general (i.e., constituting a less noisy measure), and increasingly so the more incentivized the choice was. In support of this notion, the reliability of hypothetical, non-incentive-compatible measures of preferences has been shown to be reduced (Ginon, Combris, Lohéac, Enderli, & Issanchou, 2014), and they unsystematically deviate from incentive-compatible valuation measures, such as real purchases (Cummings et al., 1995; Johannesson et al., 1998; Murphy et al., 2005). Hence, with non-incentive-compatible measures, the preference system may not always detect a relevant conflict between equally valued choice options. In consequence, this information may not be relevant enough and hence sufficient for the underlying cognitive system to demand refinement of preferences. Note that while the labeling of the endpoints of the scales differed between conditions in Experiment 1, there is no reason to assume that this affected our results. First, both were identical continuous scales, and the whole range of the scale was used in both conditions. Second, similar effects were obtained for liking scales with very heterogeneous labels (Coppin et al., 2014; Izuma et al., 2010; Koster et al., 2015; Mengarelli, Spoglianti, Avenanti, & di Pellegrino, 2013; Salti et al., 2014; Sharot et al., 2012). Third, in our between-subjects design each participant was only exposed to one or the other scale, circumventing possible anchor effects (Tversky & Kahneman, 1986) or transfer effects between scales. Finally, we empirically confirmed these results in an additional replication experiment, which used identical scale labels to the WTP experiments.

The second novel aspect of our study was that preference judgments had potential consequences. Experiment 1 alone did not provide evidence for this interpretation, as choice-induced preference change effects were observed for both items that were available and items that were not available for consumption after the experiment. However, since these items were presented in a randomized fashion to the same participants, they might have neglected the distinction between available and unavailable items to reduce cognitive effort. In this case, the existence of available items could have led to a spill-over effect on unavailable items. In the second experiment, however, two different groups of participants were presented either only with available items or only with unavailable items, and a robust choice-induced preference change effect occurred in the incentivized condition with available items only. This suggests that the effect is strongly modulated by a combination of real-world valuations and incentive compatibility of choices.

Our results suggest a model in which preferences are endogenous, that is, are a dynamic rather than a static concept that remains susceptible to changes. These findings are coherent with standard habit-formation models, which assume that preferences can change endogenously, induced by an individual's choice history (Rozen, 2010). Accordingly, past decisions may serve as a new reference point for upcoming choices, which eventually leads to the formation of a new preference (as revealed by successive choices) for the respective item. Importantly, in our study preference change was only observed after choices between equally valuable items. This may reflect the system's capacity to adapt and refine when options are concerned for which no clear preference differentials exist yet. In this case, the preference system seems to be driven to establish a hierarchy, possibly by "up-valuating" one option and/or "down-valuating" the other option. For this refinement, behavioral experiences (i.e., choosing an item, consuming it, and experiencing the consequences of consumption) would constitute an ideal driver. However, our results suggest that the choice component alone could function as a proxy for the anticipated experience, and therefore serve as "evidence" for the hedonistic value of the chosen option, triggering the observed shift in preferences. Interestingly, similar effects of previous choices on subsequent choices have been demonstrated for simple, inconsequential decisions for which no feedback was provided (Akaishi, Umeda, Nagase, & Sakai, 2014). These results suggest that in the absence of other information, the previous choice itself may function as choice-relevant evidence (Bode et al., 2012, 2014), which resembles developing a "preference" for the recently chosen choice alternative. This perspective of dynamic preference formation further aligns well with proposals that preferences may even be biased by incidental factors during choice, such as gaze and fixation duration (Krajbich, Armel, & Rangel, 2010; Shimojo, Simion, Shimojo, & Scheier, 2003).

In our study, the endogenous change in preferences could be implemented either during the choice phase (Jarcho, Berkman, & Lieberman, 2010; Kitayama et al., 2013), as an automatic by-product of the choice (Lieberman et al., 2001), or during the next valuation phase (Salti et al., 2014; Sharot et al., 2012). In a recent study, choices only biased preferences when participants spontaneously recalled or explicitly remembered their previous choices (Salti et al., 2014). This finding is consistent with cognitive dissonance theory (Festinger, 1957), which assumes that participants experience cognitive dissonance during the second valuation phase and in turn adjust current preferences to past decisions to reduce dissonance. Conceptually, another relevant classical account is self-perception theory (Bem, 1967), which states that preferences are learned by making inferences from past choices. This class of explanations, however, has been challenged by a demonstration that choice-induced preference change effects were also observed in amnestic patients, who did not have explicit memories of their past decisions (Lieberman et al., 2001). Furthermore, neuroimaging studies showed that activity in the dorsal anterior cingulate cortex (Kitayama et al., 2013) and the anterior insula (Jarcho et al., 2010; Kitayama et al., 2013) already predicted preference changes during the choice phase. These findings, however, have to be confirmed by implementing the confound controls suggested by Chen and Risen (2010). Our study was not explicitly designed to determine the origin of the observed effect but to investigate the role of incentive compatible choices. However, classical explanations referring to cognitive dissonance theory (Festinger, 1957) might predict that only preference adjustments for the rejected item should occur, while an additional devaluation of the chosen item is not necessarily required to reduce the cognitive dissonance (as choosing a preferred item does not induce cognitive dissonance). Our findings, however, clearly support the existence of a choosing bias (also see Koster et al., 2015; Sharot et al., 2012) in addition to a rejection bias. Whether another psychological mechanism or a combination of mechanisms (Lichtenstein & Slovic, 2006; see Weber & Johnson, 2009 for interdisciplinary perspectives on dynamic and context-depended preference formation) provides the best explanation at a process level could, for example, be investigated by cognitive modeling studies, which take into account differences in the margin of possible preference shifts at the upper and lower end of the valuation spectrum; or neuroimaging studies using our paradigm to investigate whether brain activity during the choice phase already predicts upcoming changes in preferences. Another interesting question arising from our work (which requires larger samples sizes) is whether individual differences in the susceptibility to choice effects exist that may be related to differences in neural activation profiles (or connectivity profiles) during choice.

In conclusion, by using willingness-to-pay for the assessment of preferences, we provided strong evidence for choice-induced preference change effects. We also showed that these effects were only robust for incentive-compatible elicitation of preferences when participants could bid for items that were available for consumption. These findings indicate, contrary to traditional accounts of decision-making (Samuelson, 1938), that preferences are not stable, but rather evolve predictably as a function of an individual's choice history (e.g., Rozen, 2010). Furthermore, they add to current existing preference theories from psychological research (Bettman et al., 1998; Fischhoff, 1991; Kahneman et al., 1982; Thaler, 1980) by indicating that new, exogenous contextual factors are not always necessary to develop our preferences; instead, preferences are susceptible to endogenous contextual factors, such as previous decisions between choice alternatives that lack strong preference differentials.

References

Akaishi, R., Umeda, K., Nagase, A., & Sakai, K. (2014). Autonomous mechanism of internal choice estimate underlies decision inertia. *Neuron, 81,* 195–206. http://dx.doi.org/10.1016/j.neuron.2013.10.018

Becker, G. M., DeGroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science, 9,* 226–232. http://dx.doi.org/10.1002/bs.3830090304

Bem, D. J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review, 74,* 183–200. http://dx.doi.org/10.1037/h0024835

Bettman, J. R., Luce, M. F., & Payne, J. W. (1998). Constructive consumer choice processes. *Journal of Consumer Research, 25,* 187–217. http://dx.doi.org/10.1086/209535

Bode, S., Murawski, C., Soon, C. S., Bode, P., Stahl, J., & Smith, P. L. (2014). Demystifying "free will": The role of contextual information and evidence accumulation for predictive brain activity. *Neuroscience and Biobehavioral Reviews, 47,* 636–645. http://dx.doi.org/10.1016/j.neubiorev.2014.10.017

Bode, S., Sewell, D. K., Lilburn, S., Forte, J. D., Smith, P. L., & Stahl, J. (2012). Predicting perceptual decision biases from early brain activity. *The Journal of Neuroscience, 32,* 12488–12498. http://dx.doi.org/10.1523/JNEUROSCI.1708-12.2012

Brehm, J. W. (1956). Postdecision changes in the desirability of alternatives. *The Journal of Abnormal and Social Psychology, 52,* 384–389. http://dx.doi.org/10.1037/h0041006

Camerer, C., & Mobbs, D. (2017). Differences in behavior and brain activity during hypothetical and real choices. *Trends in Cognitive Sciences, 21,* 46–56. http://dx.doi.org/10.1016/j.tics.2016.11.001

Chen, M. K., & Risen, J. L. (2010). How choice affects and reflects preferences: Revisiting the free-choice paradigm. *Journal of Personality and Social Psychology, 99,* 573–594. http://dx.doi.org/10.1037/a0020217

Coppin, G., Delplanque, S., Bernard, C., Cekic, S., Porcherot, C., Cayeux, I., & Sander, D. (2014). Choice both affects and reflects preferences. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 67,* 1415–1427. http://dx.doi.org/10.1080/17470218.2013.863953

Cummings, R. G., Harrison, G. W., & Rutstrom, E. (1995). Homegrown values and hypothetical surveys: Is the dichotomous choice approach incentive-compatible? *The American Economic Review, 85,* 260–266.

Egan, L. C., Bloom, P., & Santos, L. R. (2010). Choice-induced preferences in the absence of choice: Evidence from a blind two choice paradigm with young children and capuchin monkeys. *Journal of Experimental Social Psychology, 46,* 204–207. http://dx.doi.org/10.1016/j.jesp.2009.08.014

Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row, Peterso

Fischhoff, B. (1991). Value elicitation: Is there anything in there? *American Psychologist, 46,* 835–847. http://dx.doi.org/10.1037/0003-066X.46.8.835

Ginon, E., Combris, P., Lohéac, Y., Enderli, G., & Issanchou, S. (2014). What do we learn from comparing hedonic scores and willingness-to-pay data? *Food Quality and Preference, 33,* 54–63. http://dx.doi.org/10.1016/j.foodqual.2013.11.003

Izuma, K., Matsumoto, M., Murayama, K., Samejima, K., Sadato, N., & Matsumoto, K. (2010). Neural correlates of cognitive dissonance and choice-induced preference change. *Proceedings of the National Academy of Sciences of the United States of America, 107,* 22014–22019. http://dx.doi.org/10.1073/pnas.1011879108

Izuma, K., & Murayama, K. (2013). Choice-induced preference change in the free-choice paradigm: A critical methodological review. *Frontiers in Psychology, 4,* 41.

Jarcho, J. M., Berkman, E. T., & Lieberman, M. D. (2010). The neural basis of rationalization: Cognitive dissonance reduction during decision-making. *Social Cognitive and Affective Neuroscience, 6,* 460–467.

Johannesson, M., Liljas, B., & Johansson, P.-O. (1998). An experimental comparison of dichotomous choice contingent valuation questions and real purchase decisions. *Applied Economics, 30,* 643–647. http://dx.doi.org/10.1080/000368498325633

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *The American Economic Review, 93,* 1449–1475. http://dx.doi.org/10.1257/000282803322655392

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, England: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511809477

Kang, M. J., & Camerer, C. F. (2013). fMRI evidence of a hot-cold empathy gap in hypothetical and real aversive choices. *Frontiers in Neuroscience, 7,* 104.

Kitayama, S., Chua, H. F., Tompson, S., & Han, S. (2013). Neural mechanisms of dissonance: An fMRI investigation of choice justification. *NeuroImage, 69,* 206–212. http://dx.doi.org/10.1016/j.neuroimage.2012.11.034

Koster, R., Duzel, E., & Dolan, R. J. (2015). Action and valence modulate choice and choice-induced preference change. *PLoS ONE, 10,* e0119682. http://dx.doi.org/10.1371/journal.pone.0119682

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience, 13,* 1292–1298. http://dx.doi.org/10.1038/nn.2635

Lichtenstein, S., & Slovic, P. (Eds.). (2006). *The construction of preference*. Cambridge, England: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511618031

Lieberman, M. D., Ochsner, K. N., Gilbert, D. T., & Schacter, D. L. (2001). Do amnesics exhibit cognitive dissonance reduction? The role of explicit memory and attention in attitude change. *Psychological Science, 12,* 135–140. http://dx.doi.org/10.1111/1467-9280.00323

Dolan, R. J. (2010). Do decisions shape preference? Evidence from blind choice. *Psychological Science, 21,* 1231–1235. http://dx.doi.org/10.1177/0956797610379235

Lyerly, J. E., & Reeve, C. L. (2015). Development and validation of a measure of food choice values. *Appetite, 89,* 47–55. http://dx.doi.org/10.1016/j.appet.2015.01.019

Mengarelli, F., Spoglianti, S., Avenanti, A., & di Pellegrino, G. (2013). Cathodal tDCS over the left prefrontal cortex diminishes choice-induced preference change. *Cerebral Cortex, 25,* 1219–1227.

Morgenstern, R., Heldmann, M., & Vogt, B. (2014). Differences in cognitive control between real and hypothetical payoffs. *Theory and Decision, 77,* 557–582. http://dx.doi.org/10.1007/s11238-013-9408-x

Morkbak, M., Olsen, S. B., & Campbell, D. (2014). Behavioral implications of providing real incentives in stated choice experiments. *Journal of Economic Psychology, 45,* 102–116. http://dx.doi.org/10.1016/j.joep.2014.07.004

Murphy, J. J., Allen, P. G., Stevens, T. H., & Weatherhead, D. (2005). A meta-analysis of hypothetical bias in stated preference valuation. *Environmental and Resource Economics, 30,* 313–325. http://dx.doi.org/10.1007/s10640-004-3332-z

Plassmann, H., O'Doherty, J., & Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *The Journal of Neuroscience, 27,* 9984–9988. http://dx.doi.org/10.1523/JNEUROSCI.2131-07.2007

Rozen, K. (2010). Foundations of intrinsic habit formation. *Econometrica, 78,* 1341–1373. http://dx.doi.org/10.3982/ECTA7302

Salti, M., El Karoui, I., Maillet, M., & Naccache, L. (2014). Cognitive dissonance resolution is related to episodic memory. *PLoS ONE, 9*(9), e108579. http://dx.doi.org/10.1371/journal.pone.0108579

Samuelson, P. A. (1938). A note on the pure theory of consumer's behaviour. *Economica, 5,* 61–71. http://dx.doi.org/10.2307/2548836

Schonberg, T., Bakkour, A., Hover, A. M., Mumford, J. A., Nagar, L., Perez, J., & Poldrack, R. A. (2014). Changing value through cued approach: An automatic mechanism of behavior change. *Nature Neuroscience, 17,* 625–630. http://dx.doi.org/10.1038/nn.3673

Sharot, T., De Martino, B., & Dolan, R. J. (2009). How choice reveals and shapes expected hedonic outcome. *The Journal of Neuroscience, 29,* 3760–3765. http://dx.doi.org/10.1523/JNEUROSCI.4972-08.2009

Sharot, T., Fleming, S. M., Yu, X., Koster, R., & Dolan, R. J. (2012). Is choice-induced preference change long lasting? *Psychological Science, 23,* 1123–1129. http://dx.doi.org/10.1177/0956797612438733

Sharot, T., Velasquez, C. M., & Dolan, R. J. (2010). Do decisions shape preference? Evidence from blind choice. *Psychological Science, 21,* 1231–1235. http://dx.doi.org/10.1177/0956797610379235

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience, 6,* 1317–1322.

Simon, H. A. (1955). *A ural activity reveals preferences without choices* (pp. 1–47). NBER Working Paper Series.

Slovic, P. (1995). The construction of preference. *American Psychologist, 50,* 364–371. http://dx.doi.org/10.1037/0003-066X.50.5.364

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization, 1,* 39–60. http://dx.doi.org/10.1016/0167-2681(80)90051-7

Treiblmaier, H., & Filzmoser, P. (2011). *Benefits from using continuous rating scales in online survey research*. Paper presented at the International Conference on Information Systems (ICIS), Shanghai, China, December 4–7, 2011.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *The Journal of Business, 59,* S251–S278. http://dx.doi.org/10.1086/296365

Weber, E. U., & Johnson, E. J. (2009). Mindful judgment and decision making. *Annual Review of Psychology, 60,* 53–85. http://dx.doi.org/10.1146/annurev.psych.60.110707.163633

Received September 10, 2016
Revision received February 23, 2017
Accepted March 5, 2017 ■