AMERICAN PSYCHOLOGICAL ASSOCIATION
SSM Society for the Science of Motivation
Motivation Science

ISSN: 2333-8113
2023, Vol. 9, No. 4, 288–297
https://doi.org/10.1037/mot0000302

Missing Out by Pursuing Rewarding Outcomes: Why Initial Biases Can Lead to Persistent Suboptimal Choices

Chris Harris[1], Henk Aarts[1], Klaus Fiedler[2], and Ruud Custers[1]
[1] Department of Psychology, Utrecht University
[2] Department of Psychology, University of Heidelberg

While there are abundant reasons that might lead us to form wrong first impressions, further interaction (sampling) opportunities should allow us to attenuate such initial biases. Sometimes, however, theses biases persist despite repeated sampling opportunities, such as in superstitions or stereotypes. In two studies (*N*s = 100), we investigate this phenomenon. We demonstrate that in a task in which participants could repeatedly choose between two options to gain rewards, erroneous initial impressions about yielded outcomes can lead to persisting biases toward a clearly inferior option. We argue that a premature focus on reward pursuit (exploitation) rather than exploration is the cause of these biases, which persist despite plenty of opportunities and a presumed motivation to overcome them. By focusing on a supposedly best option, participants never give themselves the chance to sufficiently try out alternatives and thereby overcome their initial biases. We conclude that going for the money is not always the best strategy.

Keywords: persisting biases, sampling, exploration–exploitation tradeoff, learning by experience

Supplemental materials: https://doi.org/10.1037/mot0000302.supp

Most human behavior is goal-directed: People act to obtain outcomes they find rewarding (Custers & Aarts, 2010). This means that their behavior is for a large part based on beliefs about the relation between actions and their rewarding results. We make a joke because we believe it will cheer someone up, invite a friend over because we want to have a good time, or pick a specific restaurant because we want to have a great meal. While those beliefs may initially be based on suggestions by others (Pilditch & Custers, 2018; Pilditch et al., 2020), or other sources of knowledge, over time they increasingly become based on our first-hand experiences. At first glance, such direct experiences would seem to allow us to update our initial beliefs through repeated interaction, a feature that should protect us from relying continuously on potentially incorrect initial beliefs. We argue here, though, that this adjustment of initial beliefs is often limited when it comes to action–outcome relations exactly because people are motivated to pursue rewarding outcomes (Denrell, 2005; Rich & Gureckis, 2018; cf., Law of Effect; Thorndike, 1927). This biases their experiences as they seek out rewarding outcomes more often than nonrewarding outcomes. Hence, learning about the consequences of our behavior is inherently biased which has downstream consequences for information-integration and belief-updating processes. Here, we demonstrate one such consequence as we argue that this can even lead people to choose behaviors that yield suboptimal outcomes over other, more optimal, outcomes.

Sampling by Experience

Striving for rewarding outcomes always requires a delicate balance: On the one hand, exploiting the presumably best option may allow for maximizing the intended outcomes in the short run. On the other hand, exploring other options allows for finding and learning about potentially better alternatives. Any decision requires a choice between immediate reward pursuit of a supposedly best option or sacrificing immediate rewards to find options that will (potentially) allow for higher returns in the long run (Cohen et al., 2007; Mehlhorn et al., 2015; Mischel, 1974). Whenever a choice needs to be made, the interests of learning about alternatives and maximizing rewards are pitted directly against one another. Here, we argue that pursuit of rewarding outcomes can create an idiosyncratic subset when it comes to the outcomes people experience. That is, by exploitation of initial beliefs about actions and rewarding outcomes, people can create persisting suboptimal biases even when there is an objectively better choice alternative (Rich & Gureckis, 2018; Yechiam et al., 2001).

Of course, as long as decision makers are correct in their estimations of a given choice environment, exploitation is the optimal strategy. But when these estimations are incorrect, exploitation would lead to repeated suboptimal choices. Repeated negative outcomes of such suboptimal choices should quickly discourage further

This article was published Online First June 15, 2023.
Chris Harris https://orcid.org/0000-0001-6373-3525
All analyses and data can be found on an Open Science Framework repository https://osf.io/2z6sk/.

We thank Inge de Jong, Michelle Mojet, Julia den Ouden, and Benthe Schuurman for their help in designing and running the experiment and two anonymous reviewers for their suggestions that helped to improve the article.

The authors made the following contributions. Chris Harris contributed to conceptualization, writing–original draft preparation, and writing–review and editing. Henk Aarts contributed to writing–review and editing. Klaus Fiedler contributed to writing–review and editing. Ruud Custers contributed to conceptualization and writing–review and editing.

Correspondence concerning this article should be addressed to Chris Harris, Utrecht University, Heidelberglaan 1, 3584 CS Utrecht, The Netherlands. Email: c.a.harris@uu.nl

exploitation and instead lead to more explorative behavior. After all, who would continue betting on a losing horse? But when outcomes are repeatedly positive, a decision maker might feel confirmed in their current choice strategy and not notice that other choices would lead to even better outcomes. In other words, sufficient positive outcomes might seduce a decision maker into exploiting suboptimal choices without sufficiently exploring choice alternatives (Harris et al., 2020; Harris, Aarts, et al., 2023). It is an inherent feature of active sampling that the decision-making process constrains the information people seek and receive (Denrell, 2005; Denrell & Le Mens, 2012; Fiedler, 2000; Fiedler & Wänke, 2009; Rich & Gureckis, 2018). Betting on a winning horse can make us blind to even better alternatives.

Biases by Exploitation

In a recent line of experiments, Harris et al. (2020) and Harris, Fiedler, and Custers (2023) have demonstrated such persistent biases in reward-rich environments. Participants played two-armed bandit tasks in which they could repeatedly choose between two options that yielded positive or negative points with the total of these points later being converted to a financial reward. Following a bias induction phase, participants' choice behavior showed overall trends of exploitation across trials (Harris et al., 2020; Kasper et al., 2023) and at the individual trial level (Harris, Fiedler, & Custers, 2023), which repeatedly led to persisting biases. However, because the winning probabilities for the two choice alternatives were identical in this line of research, and any biases therefore of no actual consequence, it remains unclear whether these biases could lead people to make suboptimal choices that actually harm them.

In the current research, participants engaged in the same two-armed bandit task used by Harris et al. (2020). However, instead of two identical options, one option was objectively better than the alternative. In the induction phase, like Harris et al. (2020), we aimed to induce a bias using double-skewed distributions that are known to induce pseudocontingency illusions (Fiedler et al., 2009). Pseudocontingencies refer to a contingency heuristic that relies on base rate alignment. More specifically, because we presented one choice option more frequently to participants than the other, and because rewarding outcomes for both options occurred more frequently than losses, participants should align the (in)frequent option and the (in)frequent outcome (Fiedler & Freytag, 2004; Fiedler et al., 2013). As a result, people should show an initial preference for the frequent option, even though this option had an objectively lower probability of producing a reward than the infrequent option. In the free sampling phase that followed, frequent positive outcomes would then seemingly confirm this initial bias and result in the maintenance of this detrimental bias throughout an extended sampling period (cf., Harris et al., 2020). Because the infrequent option is actually better, we go beyond previous research and test whether participants' motivation to maximize their rewards can ironically result in foregoing better outcome probabilities.

Two things are important to highlight regarding our use of skewed distributions and hence pseudocontingencies to induce a bias. First, this particular bias induction allows us to induce a bias without manipulating the outcome probabilities. In other words, while we control the sampling frequencies of both options on the first trials, the outcomes perfectly reflect the underlying probabilities of the later phase in which participants may sample freely. Second, while we see advantages to using pseudocontingencies, we also believe that many other forms of bias induction would produce similar effects. We will return to this point in the "General Discussion." Importantly, the repeated choices we ask of participants force them into an iterative cycle: their choice behavior results in feedback which can result in updated beliefs that can inform the next choice. As long as the feedback participants receive seemingly confirms their current belief model (and the initial bias), it is extremely difficult to break out of this iterative cycle. Why should one switch horses, when clearly this one is doing well?

Probability Matching and Exploitation

What, then, constitutes a suitable benchmark in this experimental setting to which we can compare participants' behavior? Rationally speaking, the best strategy would be to exploit the higher winning probabilities of the objectively better option, as such a maximizing strategy would result in the highest payoffs. However, people rarely deploy such extreme maximizing strategy, and we therefore believe this would be too conservative a benchmark.

Instead of maximizing, people tend to match how often they choose options to the probabilities of the outcomes (probability matching; Vulkan, 2000). To the best of our knowledge, however, probability matching is not defined for choice alternatives that have outcome probabilities that are independent of one another. But, as it is assumed to emerge from a strategy of win-stay-lose-shift (WSLS; Nowak & Sigmund, 1993; Otto et al., 2011), we calculated the proportions of choices such a strategy would predict (see online supplemental material A). We therefore always first compare participants' behavior with this probability matching (PM) baseline and consider this baseline to be a reasonable approximation of what an exploitation strategy might look like in our task setting.

As a second test, we also compare participants' behavior to chance level. Choosing between both options at chance level is what we might expect if participants explore the choice alternatives randomly.

We consider both baselines to be on the conservative side. The PM baseline, because it is far from optimal exploitation in this task (maximization). The chance level baseline, because an indifference between both options would still suggest a bias in that participants clearly would not have picked up on the objectively better option. Assuming that participants learn during this task, we believe that overcoming this initial bias (e.g., through exploration) would result in participants learning that the initially infrequent choice alternative is, in fact, better and increasingly exploiting this option (e.g., similar to our PM baseline).

Experiment 1

In this first experiment, we used a distribution of evidence in which the two choice alternatives are still quite similar. While one option resulted in a positive outcome on 75% of the trials, the alternative did so on 80% of the trials.

Method

An a priori power analysis for a difference from constant *t*-test using G*Power (Faul et al., 2007) suggested a minimum sample size of at least 50 participants for an effect of choice in a condition with frequent positive outcomes for two equal options. These calculations were based on a 5% alpha-level, 80% statistical power, and effect sizes between $d = 0.35$ and $d = 0.43$ as reported by Harris et al. (2020) for maintained biases in a condition with positive outcomes on 75% of trials in Experiment 2.

A sensitivity analysis using G*Power (Faul et al., 2007) suggested that we could detect effect sizes as small as $r = .13$ ($d = 0.25$) with 80% power assuming an alpha-level of .05 and our sample size of $N = 100$.

Participants for this study were recruited via the online crowdsourcing platform Prolific Academic (https://prolific.co/) and the study was run in English on Soscisurvey (Leiner, 2020). One hundred participants ($N_{\text{female}} = 41$) with an average age of 27 years ($SD = 7.16$) participated for a financial reward of £0.85 plus additional earnings (mean £1.00, max £1.15) based on performance. All participants indicated to be fluent in English and had an approval rating of 95 (out of 100) or higher on the platform. Seventy-eight percent of participants had an educational degree of College/A levels or higher. The research line reported in this article was conducted according to the guidelines of the Ethics Review Board of the Faculty of Social and Behavioral Sciences at Utrecht University (19-155). We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the experiments (cf., Simmons et al., 2012).

Procedure

In the main task, we asked participants to make repeated choices between two bags, A and B, which resulted in a yellow or blue ball being drawn. One color would earn points, the other would lose points. Participants were instructed (and incentivized) to earn as many points as possible. The experiment consisted of four phases: An induction phase, in which participants were forced to choose particular options so that they ended up with the distribution of initial evidence that should induce a pseudocontingency (Table 1). Then, a first estimate phase followed in which participants indicated their inferences from the induction phase and which also served as a manipulation check. In a free sampling phase, participants could then choose freely between both options on each trial and earn points that would later be converted to a monetary payoff. In a final estimate phase, participants gave estimates regarding the just completed task. In total, participants completed 100 trials. We counterbalanced which bag was presented more frequently, which color was the winning color, and what color participants were asked to give estimates for. See Figure 1 for an overview over the experiment.

Induction Phase. In the induction phase, participants were first introduced to the task consisting of two bags from which they could grab either yellow or blue balls with replacement. Then, they were told that the computer had preselected which of the two bags they would be drawing from on the first few trials to get familiar with the task. On a given trial, participants would then, for example, only see bag A. After clicking on this bag and a short delay, they would receive feedback in the form of text ("You chose bag A and drew a yellow ball.") as well as an image depicting, in this case, a yellow ball. After a delay of 1 s the feedback would disappear, and the next choice was presented. Throughout the entire experiment, the current trial number and the total trial number ("Trial: *x*/100") were presented on the screen. The induction phase consisted of 17 trials in the distribution of initial evidence presented in Table 1. Importantly, while one bag was shown more frequently (12 out of 17 trials), the infrequently shown bag had the higher probability of resulting in a win (80% vs. 75% for the frequently shown bag). Thus, the co-occurrence of high base rates of winning and the frequent bag should induce a pseudocontingency illusion in favor of the frequent bag, even though the actual winning rate was higher for the other, infrequent bag.

Premeasures. Following the induction phase, we asked participants to indicate estimates regarding the distribution they had just encountered. Specifically, we asked them to first indicate a relative contingency estimate ("From which of the two bags were you more likely to grab a yellow ball?") on a slider anchored with the two bags displayed as images at the ends. We then asked participants to estimate for each bag the conditional probability of a yellow bag ("How likely was it [in %] that you grabbed a yellow ball if you chose bag A/B?") on a slider anchored at 0% and 100%. And, we asked them to indicate their confidence in both conditional probability estimates ("How confident are you that you can make a reasonable estimate regarding bag A/B?"). This phase forced participants to actively consider the evidence they had just encountered.[1] Additionally, it provided a straightforward manipulation check of the success of the bias induction.

Free Sampling. Next, we introduced the reward scheme. From this point on, participants earned 10 points for every yellow and lost 10 points for every blue ball (or vice versa). Participants were reminded that these points would be converted into a monetary reward at the end of the experiment.

In the free sampling phase, participants were free to choose either bag on each of the remaining 83 trials. Our behavioral measure was the number of choices participants made for each of the two bags during this free sampling phase. We coded every choice of the frequent bag as +1 and every choice of the alternative as 0.

Post Measures. In the final estimate phase, we asked participants to again estimate the relative contingency, conditional probabilities, and their confidence.

Data Preparation and Analyses

All dependent variables were recoded to fit a scale from [0; 1]. Accordingly, values larger than .5 for either estimate indicate a preference for the option that was shown more frequently in the induction phase. In other words, values $>.5$ would indicate a pseudocontingency bias, values $<.5$ would reflect the true underlying probabilities. For the conditional estimates, we calculated difference scores (Allan, 1980), that is, difference scores between two conditional probability estimates of yellow balls given bag A versus B, respectively, which we then recoded to a scale of [0; 1]. All comparisons are against chance level unless specified otherwise.

Data preparation and analyses were undertaken using R (Team, 2018) and in particular the packages *lme4* (Bates et al., 2015), *lmerTest* (Kuznetsova et al., 2017), and *papaja* (Aust & Barth, 2022). Across the two estimates of preference (the relative contingency estimates and the conditional probability estimates), we expected a bias toward the frequently presented option in the induction phase and therefore performed one-tailed tests. Following a successful bias induction, we expected persisting biases during sampling and again performed the same one-tailed tests on the final estimates. For the preference estimate, we used *t*-tests.

[1] Though such active consideration does not seem to be necessary (Kasper et al., 2023).

Table 1
Distribution of Initial Evidence

Choice alternative	Experiment 1 Wins	Experiment 1 Losses		Experiment 2 Wins	Experiment 2 Losses	
Frequently shown bag	9	3	75%	8	4	67%
Infrequently shown bag	4	1	80%	4	1	80%
			$\Delta p = -.05$			$\Delta p = -.13$

Note. All percentages depict the ratio of wins out of all trials for the respective location. Δp is the difference score between the conditional probabilities and describes the contingency between location and outcome (Allan, 1980).

Because the conditional estimates are bounded, we applied a logit transformation and used Wilcoxon matched-pairs signed-ranks tests. We expect no difference in participants' confidence in the two conditional estimates and performed two-tailed *t*-tests. All graphs include confidence intervals around the means, and we report confidence intervals for the effect sizes. As outlined above, we compare participants' choices and estimates to two baselines: chance level (.5) and a WSLS baseline (.444). We detail in online supplemental material A the details for obtaining the latter baseline.

In the main text, we present the central findings with regard to participants' preferences and choice behavior. We used a series of quadratic mixed-effects models to analyze participants' choices over time. Due to the binary outcomes, we fitted logistic models to the data, predicting the log odds[2] for choosing either option. The trial number was always included as a fixed effect and centered (so that the first, or for some analyses the last, free choice trial would start at 0) and scaled (for model convergence). For the end point models, we centered the models so that the last free choice trial, trial 100, would be at 0. When we included group as a factor, we allowed for random slopes for the two groups. Participants were always treated as random effects. Further details, full model specifications, and results for the models can be found in the online supplemental material B. All analyses and data can be found on an Open Science Framework repository https://osf.io/2z6sk/.

Results

Estimates: First Measurement

Following the induction phase, participants displayed a general preference toward the initially frequent but worse option. They indicated a mean preference of $M = 0.55$ ($SD = 0.33$), which differed significantly from our PM baseline $c_{\mathrm{PM}} = .444$, $t(99) = 3.35$, $p < .001$, $d = 0.33$, 95% CI [0.13, 0.54], but not from chance level, $t(99) = 1.62$, $p = .054$, $d = 0.16$, [−0.03, 0.36]. Similarly, they estimated the frequent option to be more rewarding than the actually better alternative. Specifically, the difference between their estimate scores was $\Delta p_{\mathrm{logit}} = 0.30$ ($SD = 2.18$), which again differed significantly from our PM baseline ($V = 3{,}340$, $p = .003$, $r = .28$) and also from chance level; $V = 2{,}733$, $p = .046$, $r = .16$. Finally, participants were slightly more confident in their estimate of the frequent option over that of the infrequent option, $\Delta p = .55$ ($SD = 0.13$), $t(99) = 3.83$, $p < .001$, $d = 0.38$, [0.18, 0.59].

Behavioral Data

Next, we analyzed participants' choices over time using mixed-effects models. In our first model, a significant positive intercept indicates a general bias toward the frequent (worse) choice alternative, $c = 0.59$, $z = 2.33$, $p = .020$, 95% CI [0.07; 1.09]. The significant negative estimate for trials indicates that participants' choice preference for the frequent option declined over time ($b = -1.93$, $z = -4.34$, $p < .001$) while the estimate for the squared trials indicates that this decline mainly took place on the first trials, $b = 1.57$, $z = 3.01$, $p = .003$. Finally, a nonsignificant but positive intercept at the end point of the model indicates that participants' bias was no longer different from chance, $c = 0.08$, $z = 0.29$, $p = .768$, 95% CI [−0.46; 0.57]. However, the logit of the PM baseline ($c_{\mathrm{PM}} = 0.444 \Rightarrow c_{\mathrm{logit}} = -0.22$) is not included in the 95% confidence intervals of the intercept or the end point intercept. In other words, while participants' initial bias wore off somewhat (i.e., not different from chance level), they were still far from preferring the actually better option (i.e., they did not adjust to our PM benchmark).

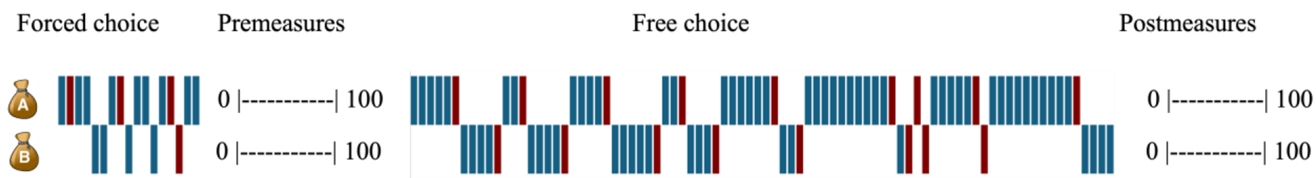
However, we were most interested in the choice behavior of participants that indicated a successful bias induction as opposed to those that did not indicate any initial bias. We therefore split participants into two groups. One, the preference-bias group ($N = 52$), contained all participants that indicated a preference for the initially frequent (worse) option on the preference measure as well as for the conditional estimates. The second, the no-preference-bias group ($N = 48$), contained all other participants. We then ran the same mixed-effects models as before but included this grouping variable.

Participants in the no-preference-bias group chose the initially infrequent (better) option more often ($c = -0.76$, $z = -2.29$, $p = .022$, 95% CI [−1.52; −0.16]) and in fact their behavior did not differ significantly from our PM baseline. Furthermore, this preference did not change over time, as indicated by the insignificant estimates for trials ($b = 0.31$, $z = 0.50$, $p = .620$) and trials squared, $b = -0.55$, $z = -0.75$, $p = .456$. At the end point, this group still displayed the same choice preference for the better option, $c = -0.88$, $z = -2.64$, $p = .008$, 95% CI [−1.52; −0.24]. This pattern differs drastically from the pattern participants in the preference-bias group displayed, $b = 2.66$, $z = 5.57$, $p < .001$. The significant interaction terms indicate that their initial preference for the frequent (worse) option ($c = 1.90$, $z = 5.53$, $p < .001$, 95% CI [1.18; 2.59]) declined over time ($b = -4.53$, $z = -5.06$, $p < .001$), but mainly in the beginning of the task, $b = 4.31$, $z = 4.10$, $p < .001$. At the end point, this group still differed significantly from the no-preference-bias group ($b = 1.84$, $z = 3.87$, $p < .001$) as well as from chance level, $c = 0.96$, $z = 2.83$, $p = .005$, 95% CI [0.28;

[2] These can be converted to probabilities using the following formula: $p = \frac{e^x}{1+e^x}$ where x is the log odd.

Figure 1

Outline of Procedure With Hypothetical Data Depicting Choices and Outcomes (Blue = Wins, Red = Losses; Lighter Gray and Darker Gray, Respectively, in the Printed Version) for the Two Choice Options A and B



Note. See the online article for the color version of this figure.

1.63]. In fact, this log odd translates to a probability of 72% for choosing the initially frequently presented (worse) option ($c_{\text{logit}} = 0.96 \Rightarrow c_{\text{prob}} = 0.72$). Given that the preference-bias group differed from chance level, they obviously also differed from our PM baseline (and accordingly the logit-value falls outside of the confidence intervals for the respective analyses). See the online supplemental materials for a more thorough discussion of the models used. Figure 2 illustrates participants' choice preferences throughout the task.

Estimates: Second Measurement

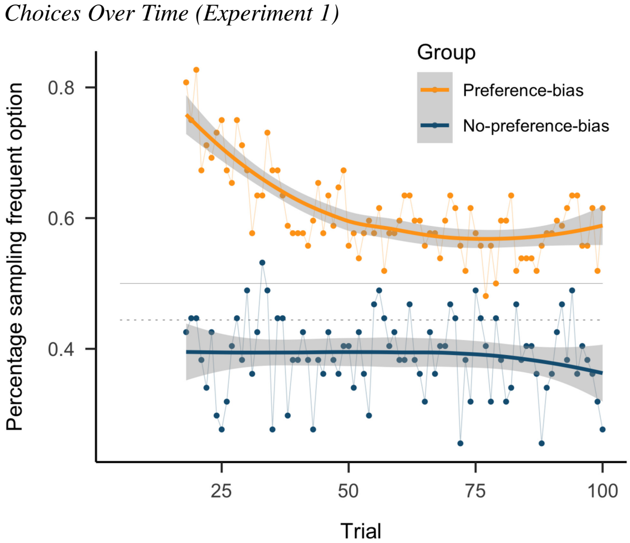
Following the free sampling phase, we asked participants the same estimates as before. In the no-preference-bias group, we see a consistent pattern of attenuation toward the better option. On the preference measure, they reported a mean of $M = 0.33$ ($SD = 0.32$), which differed from chance level ($t[47] = -3.58$, $p < .001$, $d = -0.52$, 95% CI [-0.82, -0.22]) and was even more extreme than our PM baseline, $t(47) = -2.37$, $p = .022$, $d = -0.34$, [-0.63, -0.05]. Similarly, for the contingency estimate, participants indicated a Δp_{logit} of -0.80 ($SD = 2.74$), which differed from chance level ($V = 289.5$, $p = .005$, $r = .37$) but not from the PM baseline, $V = 411$, $p = .070$, $r = .26$.

Participants in the preference-bias group, on the other hand, displayed a somewhat more complex and intriguing pattern. They reported a mean preference of $M = 0.51$ ($SD = 0.35$), which differed neither from chance level, $t(51) = 0.18$, $p = .857$, $d = 0.03$, 95% CI [-0.25, 0.30]) nor from the PM baseline, $t(51) = 1.33$, $p = .094$, $d = 0.18$, [-0.09, 0.46]. However, their conditional estimates amounted to a Δp_{logit} of 0.41 ($SD = 2.78$), which differed from chance level ($V = 814$, $p = .045$, $r = .24$) and the PM baseline, $V = 932$, $p = .014$, $r = .31$.

Participants did not differ in their confidence in their estimate of the frequent option over the infrequent option, $\Delta p = .50$ ($SD = 0.17$), $t(99) = -0.18$, $p = .855$, $d = -0.02$, 95% CI [-0.21, 0.18].

Figure 2

Choices Over Time (Experiment 1)



Note. Participants are split into two groups based on their initial preferences. The solid line is at the chance level, the dotted line at the PM baseline ($c = .444$). See the online article for the color version of this figure.

Discussion

Our manipulation successfully moved people's preferences, estimates, and initial choices away from the better option. We relied on the same method for inducing a bias as previous experiments that demonstrated persistent biases over extended sampling periods following a pseudocontingency illusion (Harris et al., 2020). The induction relied on a distribution that allowed us to differentiate between the induced bias (in which the frequent choice is paired with the frequent outcome, in this case the objectively worse option with winning) and the actual contingency (which perfectly matched the outcome probabilities throughout the free sampling phase). However, while there is support that a pseudocontingency illusion can override the actual evidence (Fiedler, 2010), it should come as no surprise that in this simple task they could also counter one another and that not all participants would fall prey to the pseudocontingency illusion.

Importantly, we replicated previous findings that initial biases can persist in reward-rich environments (Harris et al., 2020; Harris, Fiedler, & Custers, 2023). Whereas these previous studies relied on two equally rewarding choice alternatives, here we demonstrate that an initial bias can persist even in the light of an objectively better alternative. That is, after one hundred trials, participants were still more biased toward the objectively worse option than they should have been from a normative view.[3] It seems that overall frequent positive outcomes can seduce participants to skip a thorough exploration phase and exploit a supposedly better option.

Even more intriguing to us are the distinctive patterns of behavior and estimates, we find after splitting up the participants based on

[3] We ran simulations of Bayesian learning models, all of which immediately started preferring the objectively better option.

whether they showed an initial preference bias or not. The group that showed an initial preference bias for the objectively better option displayed a preference for the better choice alternative, but at the same time throughout the experiment never truly committed to a maximizing strategy (i.e., always choosing the better option). Participants that showed an initial preference for the objectively worse option, on the other hand, adjusted their initial bias somewhat but not even close to sufficiently. In fact, across almost all measures and measurement points, they did not even adjust toward chance level—though here we would consider attenuation to chance level as still being biased.

It is important to emphasize that the absence of a preference bias from chance level does not mean that participants were not attending to or properly processing the data. First of all, the significant learning curve reveals that people did adjust their behavior over time. Second, in previous studies, this manipulation produced a shift away from the objective evidence (chance) and in comparison to our PM baseline the manipulation certainly did still produce a shift. Indeed, when they could choose freely participants initially chose the worse option more often. Participants, especially in the preference-bias group, therefore, seem to have formed a preference based on the pseudocontingency illusion and after that failed to update their belief sufficiently, an effect that is known to occur in reward-rich environments (Harris et al., 2020).

Taken together, these results demonstrate that initial biases can persist even when the exploited option is actually inferior to alternatives. Even then, reward pursuit can result in behavior that is detrimental to learning. Following the initial evidence, about half of the participants preferred the objectively worse option and showed a persistent bias in the light of frequent positive outcomes. Ironically, it seems to be the case that they ended up with less reward than they could have earned exactly because they were so motivated to reap rewards.[4] In the second experiment, we increase the difference in expected value between the two choice alternatives further to investigate whether more extreme distributions will allow participants to overcome their initial biases more readily.

Experiment 2

In the second experiment, we use a more extreme distribution of evidence. Now, the initially frequently presented option results in a positive outcome in 67% of the trials whereas the alternative results in a positive outcome in 80% of the trials. This also affects the PM baseline, which is now 0.377 (again, see online supplemental material A for details).

Method

Participants for this study were again recruited via Prolific Academic https://prolific.co/ and the study was run in English on Soscisurvey (Leiner, 2020). One hundred participants ($N_{\text{female}} = 58$) with an average age of 31 years ($SD = 9.25$) participated for a financial reward of £0.85 plus additional earnings (mean £1.00, max £1.15) based on performance. All participants indicated to be fluent in English and had an approval rating of 95 (out of 100) or higher on the platform. Eighty-five percent of participants had an educational degree of College/A levels or higher. The distribution of initial evidence used in the induction phase is presented in Table 1.

Results

Estimates: First Measurement

Following the induction phase, participants displayed no general preference toward either option. They indicated a mean preference of $M = 0.49$ ($SD = 0.31$), which differed significantly from our PM baseline $c_{\text{PM}} = 0.377$, $t(99) = 3.72$, $p < .001$, $d = 0.37$, 95% CI [0.17, 0.58], but not from chance level, $t(99) = -0.20$, $p = .838$, $d = -0.02$, [−0.22, 0.18]. Similarly, they estimated both choice alternatives to be equally rewarding. Specifically, the difference between their estimate scores was $\Delta p_{\text{logit}} = 0.09$ ($SD = 1.58$), which again differed significantly from our PM baseline ($V = 3{,}516$, $p < .001$, $r = .34$) but not from chance level; $V = 2{,}363.5$, $p = .379$, $r = .03$. Participants did not differ in their confidence in their estimate of the frequent option over the infrequent option, $\Delta p = 0.51$ ($SD = 0.09$), $t(99) = 0.70$, $p = .486$, $d = 0.07$, [−0.13, 0.27].

Behavioral Data

We then analyzed participants' choices over time using the same mixed-effects models as in Experiment 1. In our first model, across all participants, the intercept lies exactly at 0 suggesting equal choice preference for the two alternatives, $c = 0.00$, $z = -0.02$, $p = .985$, 95% CI [−0.49; 0.42]. However, the PM baseline ($c_{\text{PM}} = 0.377 \Rightarrow c_{\text{logit}} = -0.50$) is not included in the 95% confidence interval. Across trials, the choice preference shifted toward preferring the initially infrequent (better) option ($b = -2.51$, $z = -5.95$, $p < .001$) but mainly on the first trials, $b = 1.96$, $z = 3.90$, $p < .001$. The negative intercept at the end point indicates a preference for the better option that roughly matches the PM baseline.

Once again, we then split up participants based on their estimates in the first part of the experiment and compared the preference-bias group ($N = 43$) with the no-preference-bias group, $N = 43$. Participants in the no-preference-bias group chose the initially infrequent (better) alternative more often ($c = -0.83$, $z = -2.75$, $p = .006$, 95% CI [−1.36; −0.24]) and did so at a rate similar to the PM baseline. Different from Experiment 1, the negative main effect for trials suggests that now participants in the no-preference-bias group increasingly chose the better alternative ($b = -2.44$, $z = -4.26$, $p < .001$) though they did so mainly in the beginning, $b = 2.42$, $z = 3.53$, $p < .001$. At the end point, they indicated a stronger preference for the better option that even differed significantly from the PM baseline, $c = -1.20$, $z = -3.95$, $p < .001$, 95% CI [−1.88; −0.65]. The preference-bias group, again differed significantly from the no-preference-bias group ($b = 1.86$, $z = 4.44$, $p < .001$) and from chance level ($c = 1.04$, $z = 3.54$, $p < .001$, 95% CI [0.45; 1.64]), but the nonsignificant interaction terms suggest that the change rate in participants' choice preference was comparable to that of the no-preference-bias group (trials: $b = -0.22$, $z = -0.26$, $p = .794$, trials^2: $b = -0.90$, $z = -0.89$, $p = .374$). At the end point, the preference-bias group seemed to have developed a slight preference for the better option, though this estimate differed neither from chance level nor the PM baseline ($c = -0.12$, $z = -0.41$, $p = .679$, 95% CI [−0.71; 0.44]) and which amounted to a probability of 47% for choosing the initially frequently presented (worse) option ($c_{\text{logit}} =$

[4] Participants in the preference-bias group earned on average £0.11 extra out of the maximum of £0.30 (36%). Participants in the no-preference-bias group earned on average £0.17 (57%) extra.

$-0.12 \Rightarrow c_{prob} = 0.47$). Figure 3 illustrates participants' choice preferences throughout the task.

Estimates: Second Measurement

Afterward, we again asked participants the same estimates as before. In the no-preference-bias group, we once again see consistent attenuation toward the better option. On the preference measure, they reported a mean of $M = 0.29$ ($SD = 0.28$), which indicates that they sampled the better option more often than chance, $t(56) = -5.75$, $p < .001$, $d = -0.76$, 95% CI [−1.06, −0.47], and even more often than our PM baseline would suggest, $t(56) = -2.41$, $p = .019$, $d = -0.32$, [−0.58, −0.05]. For the contingency estimate, participants indicated a Δp_{logit} of −0.78 ($SD = 1.72$), which differed from chance level ($V = 312$, $p < .001$, $r = .50$) and even the PM baseline, $V = 572$, $p = .044$, $r = .27$.

Participants in the preference-bias group, on the other hand, attenuated their choice preference to chance level. They reported a mean preference of $M = 0.50$ ($SD = 0.37$), which did not differ from chance level, $t(42) = -0.07$, $p = .947$, $d = -0.01$, 95% CI [−0.31, 0.29], but was higher than the PM baseline, $t(42) = 2.14$, $p = .019$, $d = 0.33$,[0.02, 0.63]. Similarly, their conditional estimates amounted to a Δp_{logit} of −0.33 ($SD = 2.26$), which did not differ from chance level ($V = 338$, $p = .835$, $r = .15$) or the PM baseline, $V = 509$, $p = .334$, $r = .07$.

Participants did not differ in their confidence in their estimate of the frequent option over the infrequent option, $\Delta p = 0.49$ ($SD = 0.13$), $t(99) = -1.16$, $p = .25$, $d = -0.12$, [−0.31, 0.08].

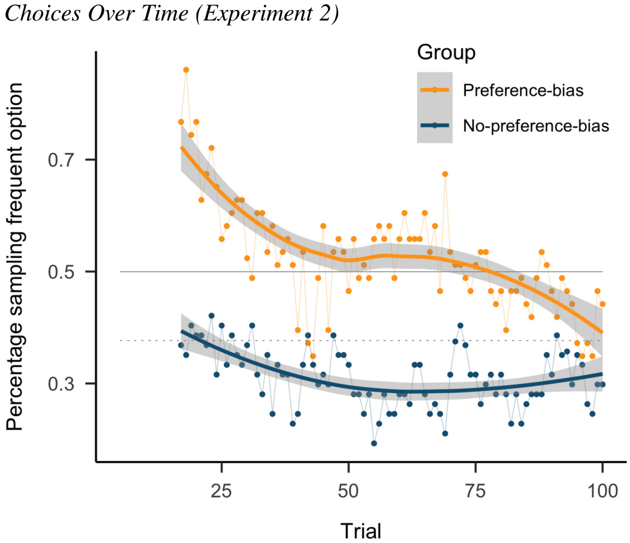
Discussion

In this second experiment, we used distributions of initial evidence that differed more extremely in their expected value than in Experiment 1. Even here, we found that initial biases persisted extraordinarily long and were only attenuated around the 75th trial. But even after the 100th trial, the estimates participants made indicate that there still remained doubt as to whether the objectively better choice alternative was indeed better. The results once again demonstrate that it was not participants' unwillingness to learn that results in the maintenance of biases. Rather, when frequent positive outcomes seduce participants to exploit one option, any alternative has to be markedly better for participants to notice.

As in any situation with two choice options in which one has better odds, maximizing the better option results in the highest payoffs (Hinson & Staddon, 1983). Even participants without an initial bias as a group did not follow this strategy, but instead displayed behavior reminiscent of PM[5] (Vulkan, 2000) and never diverged far from our PM baseline. But participants with an initial bias did not even follow this less optimal but often adaptive strategy (cf., Gaissmaier & Schooler, 2008). Instead, after the cognitive illusion seems to have overridden the genuine contingency initially, it took participants the longer part of the experiment to learn to tentatively prefer the objectively better option.[6]

Figure 3

Choices Over Time (Experiment 2)



Note. Participants are split into two groups based on their initial preferences. The solid line is at the chance level, and the dotted line at the PM baseline ($c = .377$). See the online article for the color version of this figure.

General Discussion

In two experiments, participants tried to earn as many points as possible to gain financial rewards by repeatedly choosing between two options. In an induction phase, one option was presented more frequently while the other, infrequent option resulted in a higher winning probability and was therefore the objectively better choice. About half the participants mistook the frequency of presentation to imply a contingency and incorrectly inferred that the more frequently presented option was better (Fiedler et al., 2009). These participants remained biased compared to our PM baseline (and often even to chance level) throughout the extended free choice phase and even on later estimations, thereby forfeiting better winning probabilities and ultimately financial rewards. Only participants who did not initially fall prey to this cognitive illusion correctly preferred the objectively better choice option later on.

Our results suggest that this pattern cannot be explained by apathetic participants that do not learn. After all, there clearly are changes in participants' preferences across both groups and experiments. Instead, we believe that an initial bias tempts premature exploitation and frequent wins seemingly confirm the current strategy. Why change a winning horse, as the saying goes? But by exploiting a seemingly best option, decision makers deprive themselves of opportunities to learn about choice alternatives. Intriguingly then, they end up with very idiosyncratic evidence in which one option was chosen most of the time and alternatives are never given a fair chance—people's choices lead to a highly subjective representation of the world.

Our findings demonstrate the dramatic effects this initial experience can have on later decisions. Here, we relied on skewed distributions and the pseudocontingency framework for our bias induction (Bott & Meiser, 2020; Fiedler & Freytag, 2004; Fiedler et al., 2009). We believe this framework to be suitable because information samples

[5] Note though that probability matching is not clearly defined in this case as both choice options are independent of one another. In the online supplemental material A, we offer a WSLS-approximation.

[6] Participants in the preference-bias group earned on average £0.14 extra out of the maximum of £0.30 (47%). Participants in the no-preference-bias group earned on average £0.18 (62%) extra.

in the world are not always conveniently presented at the right aggregate level (cf., illusory correlations and Simpson's paradox; Fiedler, 2000) or balanced for the decisions at hand (cf., Fiedler, 2008). Instead, certain options or outcomes may be predominant due to sampling behavior (Bott & Meiser, 2020; Denrell, 2005; Thorndike, 1927), but also simply for trivial reasons such as proximity (Back et al., 2008; Nahemow & Lawton, 1975; Preciado et al., 2012), or the potential adaptive value of relying on baserate information (Fiedler et al., 2013; Kutzner et al., 2011). What is more, from a procedural standpoint, this framework allows us to induce a bias by only manipulating the frequency in which choice options are presented on the first few trials. At the same time, however, we believe that the processes we describe here could just as well occur with different sources of initial biases. Whether it is extreme order effects (Anderson, 1965; Asch, 1946; Dennis & Ahn, 2001), communicated beliefs by others (Pilditch & Custers, 2018), or simply random fluctuation in the first few trials (Staudinger & Büchel, 2013), we assume that frequent positive outcomes reinforce an initial bias, regardless its source.

Of course, other processes (in particular motivational; Kunda, 1990) influence one's information search and belief-updating as well, and perhaps even more extremely so, which poses limitations on the generalizability of the findings and exciting avenues for further research. Nonetheless, the current article is in line with a growing literature that has put increasing emphasis on belief updating in the light of continued evidence (Alves et al., 2018; Bott & Meiser, 2020; Harris et al., 2020; Pilditch & Custers, 2018) and on the role of active sampling in decision making (Denrell & Le Mens, 2012; Fiedler & Wänke, 2009; Li et al., 2021; Prager et al., 2018; Rich & Gureckis, 2018). Decisions are rarely made in a vacuum, but instead reflect the history of the decision maker. But this history is highly subjective. It depends on earlier actions which in turn depend on beliefs. Yet, each decision, in turn, limits to what extent new information can be learned and beliefs can be updated. Learning, decision making, and the evidence we encounter together form an idiosyncratic cycle (Denrell, 2005; Harris & Custers, 2023; Li et al., 2021).

We want to highlight two important consequences of this cyclic notion of an iterative decision making and belief-updating process. First, in this particular experiment, the range of payoffs was limited and so in absolute terms the differences in earnings between the two groups might be negligible to many. But expressed in percentages, the differences already reveal that given a context with larger payoffs, the persisting bias might come costly to some. Especially in social domains, minor biases that remain uncorrected can result in long-term disadvantages, for example, for minority groups on the job market but also the examples we develop below. Moreover, in many contexts, payoffs can also grow exponentially resulting in the Matthew effect where those that already have more are also given more, thereby amplifying and further maintaining any initial differences (de Solla Price, 1965; Merton, 1968). In fact, downstream consequences of this cycle might be so impactful that some have even suggested that human's evolutionary success is (in part) due to our long childhoods offering an extended and protected exploration-encouraging period that, to some degree, circumvents the exploration–exploitation dilemma we face in adulthood (Gopnik, 2020; Liquin & Gopnik, 2022).

Second, this notion of subjective experience can explain a wide range of social phenomena. Why do some people believe so strongly in the effectiveness of certain alternative medicines while the medical sciences at best find mixed evidence? One's subjective experience might suggest a strong contingency: When I have the flue, I take my remedy, and within a few days, I feel better. People want to get healthy quickly, and so they rely on what seems to have worked in the past. Few people, when sick, are willing to engage in careful hypothesis testing. But not only are some people using less effective treatments, the costs of such behavior might be considerable for health care systems. Perhaps even more striking are the effects such persisting biases can have in social interactions. Initial biases about, for example, the trustworthiness of interaction partners will lead to either engagement or to their rejection and affect to what extent biases can be updated or remain unchanged with detrimental consequences for interaction partners. By being shunned, they never receive the opportunity to disprove the existing stereotypes and remain systematically disadvantaged (Fetchenhauer & Dunning, 2010; Jaeger et al., 2023). The current research demonstrates that biases may persist even when they are of disadvantage to the individuals holding the biases.

As such, these processes might even be at the heart of stereotype maintenance. Existing stereotypes will factor in when deciding whom to interact with in social settings. After all, when seeking rewarding interactions, one would rely on their beliefs about and impressions of others. And as long as one encounters sufficient rewarding interactions in the majority group, there might not be any reason to doubt one's strategy. As a consequence, however, instead of reducing the discrepancy in information and interaction with minority group members that often leads to stereotypes in the first place, this imbalance is maintained or even strengthened (Alves et al., 2018; Denrell & Le Mens, 2011; Kutzner & Fiedler, 2017). It comes as no surprise that interventions often focus on increasing contact between majority and minority groups (Pettigrew & Tropp, 2006). But how can stereotypes persist also in situations in which there are opportunities for contact between majority and minority groups? The current research suggests that we might not be taking advantage of such opportunities because we focus too strongly on the supposedly best choice alternatives, the majority group we know well, and neglect alternatives, such as less known minority groups.

In conclusion, one's information-sampling strategies induced by one's early experience can markedly constrain the extent to which beliefs can ever be updated. This can lead to discrepancies between one's subjective experience and the objective world, the consequences of which can lead to beliefs and choices that are harmful to the individual, interaction partners, and society as a whole.

References

Allan, L. G. (1980). A note on measurement of contingency between two binary variables in judgment tasks. *Bulletin of the Psychonomic Society*, *15*(3), 147–149. https://doi.org/10.3758/BF03334492

Alves, H., Koch, A., & Unkelbach, C. (2018). A cognitive-ecological explanation of intergroup biases. *Psychological Science*, *29*(7), 1126–1133. https://doi.org/10.1177/0956797618756862

Anderson, N. H. (1965). Primacy effects in personality impression formation using a generalized order effect paradigm. *Journal of Personality and Social Psychology*, *2*(1), 1–9. https://doi.org/10.1037/h0021966

Asch, S. E. (1946). Forming impressions of personality. *The Journal of Abnormal and Social Psychology*, *41*(3), 258–290. https://doi.org/10.1037/h0055756

Aust, F., & Barth, M. (2022). *papaja: Prepare reproducible APA journal articles with R Markdown*. https://github.com/crsh/papaja

Back, M. D., Schmukle, S. C., & Egloff, B. (2008). Becoming friends by chance. *Psychological Science*, *19*(5), 439–440. https://doi.org/10.1111/j.1467-9280.2008.02106.x

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

Bott, F. M., & Meiser, T. (2020). Pseudocontingency inference and choice: The role of information sampling. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(9), 1624–1644. https://doi.org/10.1037/xlm0000840

Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1481), 933–942. https://doi.org/10.1098/rstb.2007.2098

Custers, R., & Aarts, H. (2010). The unconscious will: How the pursuit of goals operates outside of conscious awareness. *Science*, *329*(5987), 47–50. https://doi.org/10.1126/science.1188595

Dennis, M. J., & Ahn, W.-K. (2001). Primacy in causal strength judgments: The effect of initial evidence for generative versus inhibitory relationships. *Memory & Cognition*, *29*(1), 152–164. https://doi.org/10.3758/BF03195749

Denrell, J. (2005). Why most people disapprove of me: Experience sampling in impression formation. *Psychological Review*, *112*(4), 951–978. https://doi.org/10.1037/0033-295X.112.4.951

Denrell, J., & Le Mens, G. (2011). Seeking positive experiences can produce illusory correlations. *Cognition*, *119*(3), 313–324. https://doi.org/10.1016/j.cognition.2011.01.007

Denrell, J., & Le Mens, G. (2012). Social judgments from adaptive samples. In J. I. Krueger (Ed.), *Social judgment and decision making* (pp. 151–169). Psychology Press.

de Solla Price, D. J. (1965). Networks of scientific papers. *Science*, *149*(3683), 510–515. https://doi.org/10.1126/science.149.3683.510

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fetchenhauer, D., & Dunning, D. (2010). Why so cynical? Asymmetric feedback underlies misguided skepticism regarding the trustworthiness of others. *Psychological Science*, *21*(2), 189–193. https://doi.org/10.1177/0956797609358586

Fiedler, K. (2000). Beware of samples! A cognitive-ecological sampling approach to judgment biases. *Psychological Review*, *107*(4), 659–676. https://doi.org/10.1037/0033-295X.107.4.659

Fiedler, K. (2008). The ultimate sampling dilemma in experience-based decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(1), 186–203. https://doi.org/10.1037/0278-7393.34.1.186

Fiedler, K. (2010). Pseudocontingencies can override genuine contingencies between multiple cues. *Psychonomic Bulletin & Review*, *17*(4), 504–509. https://doi.org/10.3758/PBR.17.4.504

Fiedler, K., & Freytag, P. (2004). Pseudocontingencies. *Journal of Personality and Social Psychology*, *87*(4), 453–467. https://doi.org/10.1037/0022-3514.87.4.453

Fiedler, K., Freytag, P., & Meiser, T. (2009). Pseudocontingencies: An integrative account of an intriguing cognitive illusion. *Psychological Review*, *116*(1), 187–206. https://doi.org/10.1037/a0014480

Fiedler, K., Kutzner, F., & Vogel, T. (2013). Pseudocontingencies: Logically unwarranted but smart inferences. *Current Directions in Psychological Science*, *22*(4), 324–329. https://doi.org/10.1177/0963721413480171

Fiedler, K., & Wänke, M. (2009). The cognitive-ecological approach to rationality in social psychology. *Social Cognition*, *27*(5), 699–732. https://doi.org/10.1521/soco.2009.27.5.699

Gaissmaier, W., & Schooler, L. J. (2008). The smart potential behind probability matching. *Cognition*, *109*(3), 416–422. https://doi.org/10.1016/j.cognition.2008.09.007

Gopnik, A. (2020). Childhood as a solution to explore–exploit tensions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1803), Article 20190502. https://doi.org/10.1098/rstb.2019.0502

Harris, C., Aarts, H., Fiedler, K., & Custers, R. (2023). *Iterative cycles in psychology: How hedonic outcomes, idiosyncratic experiences, and decisions by experience shape our subjective world*. [Manuscript in Preparation]. Department of Psychology, Utrecht University.

Harris, C., & Custers, R. (2023). Biased preferences through exploitation. In K. Fiedler, P. Juslin, & J. Denrell (Eds.), *Sampling in judgement and decision making* (pp. 207-221). Cambridge University Press.

Harris, C., Fiedler, K., & Custers, R. (2023). *When biases beget biases: Erroneous first impressions of choice options thrive in settings with graded outcomes*. [Manuscript Submitted for Publication]. Department of Psychology, Utrecht University.

Harris, C., Fiedler, K., Marien, H., & Custers, R. (2020). Biased preferences through exploitation: How initial biases are consolidated in reward-rich environments. *Journal of Experimental Psychology: General*, *149*(10), 1855–1877. https://doi.org/10.1037/xge0000754

Hinson, J. M., & Staddon, J. E. R. (1983). Matching, maximizing, and hill-climbing. *Journal of the Experimental Analysis of Behavior*, *40*(3), 321–331. https://doi.org/10.1901/jeab.1983.40-321

Jaeger, B., Harris, C., Custers, R., & Todorov, A. (2023). *Biased partner selection propagates stereotypes and inequalities in social exchange*. [Manuscript Submitted for Publication]. Department of Social Psychology, Tilburg University.

Kasper, J., Fiedler, K., Kutzner, F., & Harris, C. (2023). On the role of exploitation and exploration strategies in the maintenance of cognitive biases: Beyond the pursuit of instrumental rewards. *Memory & Cognition*, 1–14. https://doi.org/10.3758/s13421-023-01393-8

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480–498. https://doi.org/10.1037/0033-2909.108.3.480

Kutzner, F., & Fiedler, K. (2017). Stereotypes as pseudocontingencies. *European Review of Social Psychology*, *28*(1), 1–49. https://doi.org/10.1080/10463283.2016.1260238

Kutzner, F., Vogel, T., Freytag, P., & Fiedler, K. (2011). Contingency inferences driven by base rates: Valid by sampling. *Judgment and Decision Making*, *6*(3), 211–221. https://doi.org/10.1017/S1930297500001418

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). Lmertest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Leiner, D. J. (2020). *SoSci Survey*. (Version 3.4.01) [Computer software]. Available at https://www.soscisurvey.de

Li, A. X., Gureckis, T. M., & Hayes, B. K. (2021). Can losses help attenuate learning traps? *Proceedings of the Annual Meeting of the Cognitive Science Society*, Vol. 43. https://escholarship.org/uc/item/85z4p46m

Liquin, E. G., & Gopnik, A. (2022). Children are more exploratory and learn more than adults in an approach-avoid task. *Cognition*, *218*, Article 104940. https://doi.org/10.1016/j.cognition.2021.104940

Mehlhorn, K., Newell, B. R., Todd, P. M., Lee, M. D., Morgan, K., Braithwaite, V. A., Hausmann, D., Fiedler, K., & Gonzalez, C. (2015). Unpacking the exploration–exploitation tradeoff: A synthesis of human and animal literatures. *Decision*, *2*(3), 191–215. https://doi.org/10.1037/dec0000033

Merton, R. K. (1968). The Matthew effect in science: The reward and communication systems of science are considered. *Science*, *159*(3810), 56–63. https://doi.org/10.1126/science.159.3810.56

Mischel, W. (1974). Processes in delay of gratification. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 7, pp. 249–292). Academic Press.

Nahemow, L., & Lawton, M. P. (1975). Similarity and ropinquity in friendship formation. *Journal of Personality and Social Psychology*, *32*(2), 205–213. https://doi.org/10.1037/0022-3514.32.2.205

Nowak, M., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature*, *364*(6432), 56–58. https://doi.org/10.1038/364056a0

Otto, A. R., Taylor, E. G., & Markman, A. B. (2011). There are at least two kinds of probability matching: Evidence from a secondary task. *Cognition*, *118*(2), 274–279. https://doi.org/10.1016/j.cognition.2010.11.009

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, *90*(5), 751–783. https://doi.org/10.1037/0022-3514.90.5.751

Pilditch, T. D., & Custers, R. (2018). Communicated beliefs about action-outcomes: The role of initial confirmation in the adoption and maintenance of unsupported beliefs. *Acta Psychologica*, *184*, 46–63. https://doi.org/10.1016/j.actpsy.2017.04.006

Pilditch, T. D., Madsen, J. K., & Custers, R. (2020). False prophets and Cassandra's curse: The role of credibility in belief updating. *Acta Psychologica*, *202*, Article 102956. https://doi.org/10.1016/j.actpsy.2019.102956

Prager, J., Krueger, J. I., & Fiedler, K. (2018). Towards a deeper understanding of impression formation—New insights gained from a cognitive-ecological perspective. *Journal of Personality and Social Psychology*, *115*(3), 379–397. https://doi.org/10.1037/pspa0000123

Preciado, P., Snijders, T. A., Burk, W. J., Stattin, H., & Kerr, M. (2012). Does proximity matter? Distance dependence of adolescent friendships. *Social Networks*, *34*(1), 18–31. https://doi.org/10.1016/j.socnet.2011.01.002

Rich, A. S., & Gureckis, T. M. (2018). The limits of learning: Exploration, generalization, and the development of learning traps. *Journal of Experimental Psychology: General*, *147*(11), 1553–1570. https://doi.org/10.1037/xge0000466

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012, October 14). A 21 word solution. *SSRN Electronic Journal*. http://dx.doi.org/10.2139/ssrn.2160588

Staudinger, M. R., & Büchel, C. (2013). How initial confirmatory experience potentiates the detrimental influence of bad advice. *Neuroimage*, *76*, 125–133. https://doi.org/10.1016/j.neuroimage.2013.02.074

Team, R. C. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Thorndike, E. L. (1927). The law of effect. *The American Journal of Psychology*, *39*(1/4), 212–222. https://doi.org/10.2307/1415413

Vulkan, N. (2000). An economist's perspective on probability matching. *Journal of Economic Surveys*, *14*(1), 101–118. https://doi.org/10.1111/1467-6419.00106

Yechiam, E., Erev, I., & Gopher, D. (2001). On the potential value and limitations of emphasis change and other exploration-enhancing training methods. *Journal of Experimental Psychology: Applied*, *7*(4), 277–285. https://doi.org/10.1037/1076-898X.7.4.277

Received November 23, 2022
Revision received May 10, 2023
Accepted May 13, 2023 ■