Too Good to Be True? Psychological Responses to Surprising Options in Risk–Reward Environments

Christina Leuker, Thorsten Pachur, Ralph Hertwig, and Timothy J. Pleskac

Center for Adaptive Rationality,

Max Planck Institute for Human Development,

Berlin, Germany

24 January, 2018

Word count: 7881

Author Note

Christina Leuker, Thorsten Pachur, Ralph Hertwig, and Timothy J. Pleskac, Center for Adaptive Rationality, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany.

Author contributions were as follows: Conceptualization: C.L., T.P., R.H., & T.J.P.; Methodology: C.L., T.P. & T.J.P.; Software: C.L.; Data collection and curation: C.L.; Formal analysis: C.L.; Writing original draft: C.L.; Reviewing & Editing: C.L., R.H., T.P., & T.J.P.; Supervision: T.J.P.

The first author was supported by the MaxNetAging Research School at the Max Planck Institute for Demographic Research in Rostock, Germany. We would like to thank Chantal Wysocki and Lisa-Marie Jirak for assistance with data collection, and Susannah Goss for editing the manuscript. All gamble stimuli, data and analyses can be retrieved via the Open Science Framework (anonymous link for review).

Correspondence concerning this article should be addressed to Christina Leuker, Max Planck Institute for Human Development, Center for Adaptive Rationality, Lentzeallee 94, 14195 Berlin, Germany. E-mail: leuker@mpib-berlin.mpg.de

Abstract

Options that sound too good to be true often *are* too good to be true. How can decision makers detect if something is too good to be true, and what are their psychological responses to such options? Here, we argue that they know about and compare them against a regularity present in many domains in the environment: Risks and rewards are typically inversely related (Pleskac & Hertwig, 2014). Using a novel preferential oddball paradigm, we investigated behavioral and physiological responses to deviations from such regularities as participants evaluated risky options. In two experiments ($N = 183$), participants priced monetary gambles drawn from environments in which risks and rewards were negatively correlated, positively correlated, or uncorrelated. In later trials, they were presented with "surprising" gambles that deviated from the respective environment's risk–reward structure. Pricing, response times and (in Experiment 2) pupil dilation were recorded. In both experiments, participants took more time responding to "surprising" than to "expected" options. Crucially, "surprisingly good" options were evaluated with even more scrutiny than "surprisingly bad" options. When response time was limited, prices for surprising options offering high payoffs deviated from the expected values of the gambles in a way that was consistent with a distortion towards the environment's risk–reward structure. Moreover, surprisingly good options were associated with an increase in pupil size. Taken together, risky options are evaluated based not only on their payoffs and probabilities (evaluation from givens), but also on the extent to which they fit the risk–reward structure of the environment (evaluation from environments).

Too Good to Be True? Psychological Responses to Surprising Options in Risk–Reward Environments

On December 25th, 2017, Nicole Coggins and her mother-in-law bought 100 lottery tickets for the South Carolina Holiday Cash Add-a-Play lottery for $1 each. Astonishingly, every ticket was a winner, giving them a total of $18,000 in winnings. Wade Crenshaw, who was working behind the cash register at a convenience store that day, noticed that more and more people were asking to buy Add-A-Play tickets. He later said, "It was weird, everybody [was] winning so much. I didn't know if they were doing some kind of Christmas special" (Fortin, 2017). If you also think that this sounds "too good to be true," you'd be right: When the lucky winners went to cash in their prizes, the machine deemed their tickets to be invalid (Harrington, 2017). As it later emerged, the many winning lottery tickets were not a Christmas special, but could be traced back to a computer glitch.[1]

As this story shows, if something sounds too good to be true, it probably is. But how do people tell that something is too good to be true? The only way to detect such exceptionally good gambles is by being aware of the tradeoffs one faces in *regular*, or nonexceptional, choice situations. A frequent tradeoff is the one between payoffs and probabilities: Safer gambles (e.g., bonds) usually yield smaller payoffs than riskier gambles (e.g., stocks). Higher payoffs are more unlikely in the lottery and many other monetary and nonmonetary domains (Pleskac & Hertwig, 2014). In other words, risks and rewards are inversely related in many of life's gambles. In this article, we examine to what extent people adapt to regularities in risk–reward relationships, and how learning and adapting to different risk–reward structures influences their responses to options that surprise them by *deviating* from the learned regularity. Specifically, we were interested in behavioral and physiological responses to these surprising options.

[1] As of January 23, 2018, the South Carolina Education Lottery Commission may, pending an internal investigation, treat the the tickets as valid and owe more than $35 million to all the winners from December 25, 2017 ("South Carolina may owe double in Christmas lottery glitch", 2018).

If people showed specific responses to such surprising options, it would indicate that they do not assess the options' values soley in terms of their payoffs and risks, which we term "evaluation from givens." Rather, it would indicate they also assess how the current option relates to the larger population of options, which we term "evaluation from environments." By and large, theories of decision making are mute with regard to how people take into account the specific choice environment (though see Payne et al., 1993; Todd et al., 2012). Instead, they focus on how people form a preference among the options available in the choice set (e.g., Birnbaum, 1992; Busemeyer & Diederich, 2002; Kahneman & Tversky, 1979; Mellers et al., 1997; Trueblood et al., 2014; Usher & McClelland, 2004). An increasing amount of evidence suggests, however, that how options are evaluated depends in part on the ecological (marginal) distribution of monetary payoffs and probabilities (Birnbaum, 1992; Stewart et al., 2006, 2015; Walasek & Stewart, 2015). As mentioned before, here, we highlight an additional ecological regularity that focuses on the *joint* distribution of payoffs and probabilities that make up the options. While risks and rewards (or payoffs and probabilities) generally tend to be inversely related, the strength of these risk–reward relationships varies across domains: For gambles in casinos, for instance, there is a near-perfect (though biased) inverse relationship between payoffs and probabilities; other gambles, such as which journal to submit a paper to (trading off impact factor against acceptance rate), show a decidedly weaker relationship (Pleskac & Hertwig, 2014). Moreover, a risk–reward relationship is not always a given. For instance, a negative relationship between risk and reward is to be expected in economic markets only if an equilibrium has been achieved, but not in newly forming markets that have not yet reached an equilibrium.

Our previous work has shown that people indeed learn such risk–reward structures present in their immediate choice environment (Leuker et al., 2017a). Specifically, we asked people to evaluate monetary gambles of the form "p chance of winning x, otherwise nothing" by indicating a price for each gamble (where p was the probability and x was the monetary payoff). Participants were exposed to environments in which risks and rewards were negatively correlated, positively correlated, or uncorrelated. They were not informed about the different risk–reward structures, nor were they instructed to learn the relationship. Thus, any learning was unsupervised (where there is no corrective feedback; Love, 2002) and incidental (where learning is not the primary objective; Brooks, 1978; Dulany et al., 1984; Nelson, 1984; Ward & Scott, 1987; Wattenmaker, 1991; Whittlesea, 1987). When they were asked to estimate the probability of winning different payoffs, the participants' estimates reflected the environment to which they had been exposed. Moreover, they used the learned risk–reward relationship to make decisions under uncertainty, in which the probabilities of the possible outcomes were unknown. When payoffs were large, participants who had previously experienced a "negative risk–reward world" preferred a smaller, sure reward over a larger, uncertain reward. This pattern is consistent with participants inferring low probabilities for high payoffs. In contrast, participants in the "positive risk–reward world" preferred the larger, uncertain option over the smaller, sure thing, consistent with inferring high probabilities for high rewards.

Whereas our previous work investigated how people exploit learned risk–reward structures in decisions under uncertainty, here we address whether—and if so to what extent—risk–reward structures also leak into the evaluation of risky gambles in which both payoffs and probabilities are explicitly given (decisions under risk). Here, people have all the information necessary to choose in a way that is consistent with maximizing subjective expected utility (Savage, 1954; von Neumann & Morgenstern, 1944). Yet regularities in the combination of risks and rewards in an environment may also lead people to form expectations about the risks and rewards offered by an option. Options that do not match the environmental structure may be perceived as "surprising" and consequently trigger specific behavioral and physiological responses. The goal of this article is to identify and characterize those responses.

Responses to surprising stimuli have been studied in other areas of psychology and neuroscience. Researchers using what has become known as the oddball paradigm (Herrmann & Knight, 2001; Huettel & McCarthy, 2004; Picton, 1992; Squires et al., 1975) have examined whether participants automatically detect surprising stimuli in a sequence (called "deviants" or "oddballs"), without being explicitly instructed to process a particular sequence or deviations from it. For example, in a typical auditory oddball paradigm, participants are presented with a sequence of standards (s) and a small subset of deviants (d) that can differ in volume, duration, or pitch (e.g., s–s–s–s–s–s–d–s–s–d–s–s–s). The deviants are detected even when participants' attention is not focused on the stimuli. The detection of deviants typically has an electrophysiological correlate, the mismatch negativity (for reviews, see Näätänen, 2007; Garrido et al., 2009).

Do oddball effects also occur when people evaluate a sequence of options in preferential, risky choice, and encounter an option that deviates from the learned regularities in the combination of payoff and probability? How do people respond to such "risk–reward oddballs"? Unlike in the perceptual studies in which oddballs have previously been observed, the stimuli used in decisions under risk are rather abstract, higher level objects. Moreover, in risk–reward environments, *different* options drawn from one and the same environment can be consistent with the same risk–reward environment (e.g., \$10 with $p = .9$; \$20 with $p = .8$; and \$30 with $p = .7$ all fit a negative risk–reward environment; see Leuker et al., 2017a), whereas the standards in an oddball paradigm are identical across the sequence (s–s–s–s–s). Finally, other than in the perceptual studies, gambles are *value–based* stimuli. As a consequence, deviants from a risk–reward structure can either be "surprisingly good" (higher in expected value than expected) or "surprisingly bad" (lower in expected value than expected). Note that in this article we are concerned with peoples' evaluations of options in the gain domain; thus, "surprisingly bad" options still involve a gain (but a small and unlikely one).

Physiological responses to "surprising" stimuli have been investigated in reward-based paradigms, where it has been shown that both human and animals are sensitive to "risk-prediction errors"—that is, to mismatch between expected and experienced risks, or probabilities (Bossaerts, 2010; O'Neill & Schultz, 2013; Preuschoff et al., 2008). One key finding is that mispredictions in either direction (probability is low but a reward is still obtained, or probability is high but no reward is obtained; see Preuschoff et al., 2011) seem to be associated with pupil dilation.[2] In our case, the nature of surprise is somewhat different, as it is based not on the outcome, but on the mismatch between the properties of the current option (probabilities and payoffs) and its environment. In addition, detecting surprising options in risk–reward environments does not necessitate feedback about whether or not an option is obtained. We return to these conceptual (and possibly functional) differences between our study and previous studies of surprise in the Discussion.

The Present Experiments

In two experiments, we investigated behavioral (Experiments 1 & 2) and physiological (Experiment 2) responses to surprising (expected) options that deviated from (vs. matched) a previously learned risk–reward environment. Between participants, the structure of the environment was manipulated to be negatively correlated, positively correlated, or uncorrelated (Figure 1). In both experiments, participants indicated the price at which they would be willing to sell a monetary gamble of the form "p chance of winning x, otherwise nothing." There were two types of gambles. *Environment gambles* defined the structure of the environment in a given condition. *Test gambles* were common to all conditions but, depending on the condition, were consistent with the environmental structure (expected gambles), inconsistent with the environmental structure (surprising gambles), or (in the uncorrelated condition) neither to be expected nor perceived as

[2] Reward-prediction errors, in contrast, involve a mismatch between expected and experienced rewards. In contrast to risk-prediction errors, the neurobiological correlates differ depending on the *direction* of the misprediction: Rewards better than expected lead to a positive reward-prediction error and an increase in dopaminergic firing; rewards worse than expected lead to a negative reward-prediction error and a decrease in dopaminergic firing (Schultz, 2002; Schultz et al., 1997).

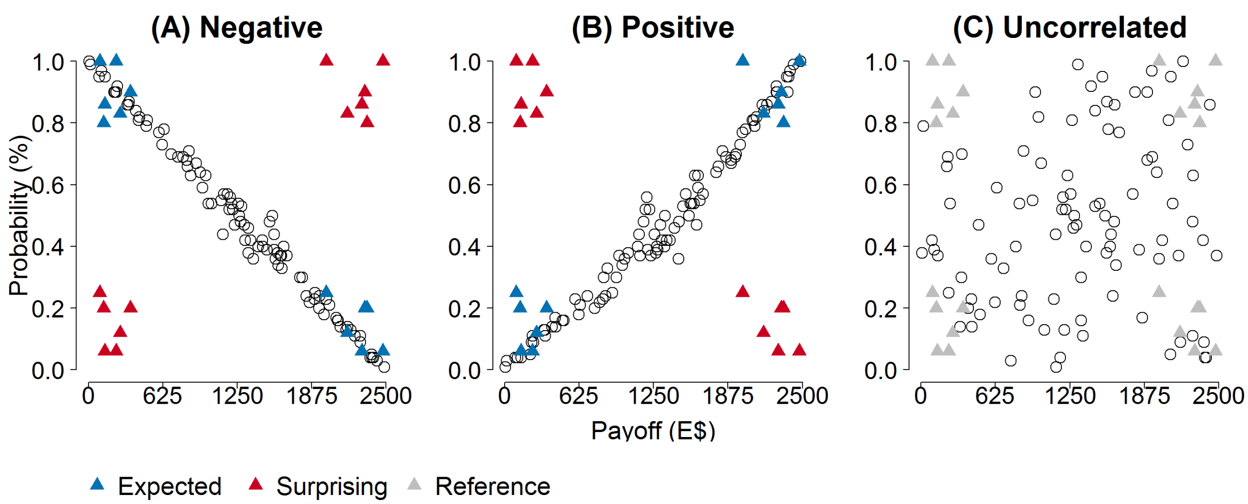

Figure 1. Stimuli. Gambles were drawn from one of the three risk–reward environments. A set of test gambles, which was common to all three conditions, was randomly interspersed after two-thirds (Experiment 1) or one-third (Experiment 2) of the trials. The number and proportions of environment gambles vs. test gambles varied slightly between experiments. Test gambles are color-coded by gamble type: Gambles shown in blue were consistent with the risk–reward relationships in a condition and could therefore be expected. Gambles shown in red were inconsistent with risk–reward relationships in a condition and were therefore surprising. Gambles shown in light gray (panel C) served as reference gambles. Here, participants were unlikely to have any expectations about particular risk–reward relationships (due to the environment gambles being uncorrelated).

surprising (reference gambles). After some exposure to one experimental environment—through evaluating a number of environment gambles—participants were presented with the test gambles.

Participants were not instructed to pay attention to the underlying risk–reward structure in either experiment. We examined how pricing and response times (RTs) varied as a function of whether the gambles were surprising or expected, and relative to the reference gambles in the uncorrelated condition. In Experiment 2, in addition to the behavioral responses, we tracked pupil size in response to surprising, expected, and reference gambles. Next, we formulate our hypotheses for the various dependent variables in our experiments.

Pricing

In both experiments, participants' pricing decisions were incentivized using a Becker-DeGroot-Marschak auction (Becker et al., 1964). The prices stated for surprising options could deviate from the expected value in various ways: On the one hand, participants could rely solely or partially on the risk–reward structure they have experienced to price the options. For instance, if a surprising gamble is associated with a high probability but the risk–reward structure suggests a low probability, participants might indicate a price that is lower than the gamble's expected value. Conversely, if a surprising gamble is associated with a low probability but the risk–reward structure suggests a high probability, participants might indicate a price that is higher than the gamble's expected value.[3]

Response times

Increasing familiarity with the risk–reward structure of an environment may accelerate the processing of subsequent options consistent with that structure. At the same time, people may need more time to price surprising gambles (e.g., to integrate the payoff with the "surprising probability")—than when the same gambles are presented in a context in which their structure is expected. This pattern would be consistent with findings in the domain of event sequence learning (i.e., responses to stimuli in different locations that follow a sequence, e.g., 4–3–2–1–4–3–2–1). Here, longer RTs for stimuli inconsistent with a

[3] A qualitatively identical prediction is that (some) people align payoff information to probabilities (instead of probability information to payoffs): If a surprising gamble is associated with a high payoff, but the risk–reward structure suggests a low payoff, participants would be expected to indicate a lower price than warranted by the gamble's expected value.

learned sequence are taken as direct evidence that a sequential stimulus structure has been learned (Rüsseler & Rösler, 2000).

As noted earlier, a unique feature of surprising value-based stimuli is that options can be surprisingly good or surprisingly bad. Prior research found that people may have a mechanism in place that "prevent[s] impulsive responding due to the presence of high value options" (Cavanagh et al., 2014, p. 2) (also see Frank et al., 2007). By extension, participants may respond more slowly to surprisingly good options than to surprisingly bad ones. Note such a mechanism would also be consistent with response times increasing not as a function of surprise (in evaluations from environments), but as a function of absolute value (in evaluations from givens).

Pupil dilation

In Experiment 2, we modified the pricing paradigm used in Experiment 1 to measure pupil dilation as participants inspected the properties of a gamble. We did so by sequentially presenting first the payoff, then the probability for each gamble. We predicted that participants would be surprised by options for which the probability information deviated from the learned risk–reward environment, and that their surprise would become manifest as the probability information was revealed. Inspired by paradigms investigating feedback-based "risk-prediction errors" (Preuschoff et al., 2011), we hypothesized that participants would show greater pupil dilation when an option turned out to be surprising (i.e., to have an unexpected probability), but not in response to just seeing high or low payoffs.

Pupil dilation may also be linked to surprising options for different reasons. Recent research has shown that pupil dilation is associated with an increase in the decision threshold in decisions between very similar options (see Cavanagh et al., 2014, for choices and pupil dilation modeled as a drift diffusion process). This decision threshold results in a more rigorous evaluation of the alternatives, which in turn produces longer RTs. A similar mechanism might emerge for surprising options in risk–reward environments, in which a more rigorous evaluation of surprising options may be linked to both longer RTs and increases in pupil size. Here, the effect of surprise on pupil dilation could also depend on the option being “surprisingly good” versus “surprisingly bad” or on the payoffs offered by the surprising option. If RTs indicate a more rigorous evaluation of surprisingly good (relative to surprisingly bad) options, a similar dissociation may be observed in pupillary signals. Thus, it is plausible that participants scrutinize more carefully only those surprising options that offer a surprisingly high expected value and not those that offer a surprisingly low one (e.g., due to less impulsive responding in the presence of high-value options; Frank et al., 2007).

Experiment 1: Behavioral Responses To Surprising Risk–Reward Combinations

In Experiment 1, participants priced monetary gambles drawn from negative, positive, or uncorrelated risk–reward environments. Our main question was how RTs and prices would differ when participants encountered gambles that represented surprising risk–reward combinations relative to expected risk–reward combinations (in the positive and negative environments) and relative to reference gambles (i.e., in the uncorrelated environment, when gambles could not be compared against any risk–reward regularity). We should note that we originally designed Experiment 1 as a means of testing how exposure to different risk–reward environments impacted performance on various *subsequent* decision-making tasks. These data and analyses are reported elsewhere (Experiment 2 in Leuker et al., 2017a).

Method

Participants. We recruited 90 participants (53 females, age 24.7 years, $SD = 4.1$ years, proportion students = .72) from the participant pool maintained at the Max Planck Institute for Human Development. Each participant completed the experiment (duration 65 min) in exchange for a show-up fee of €10 and a performance-based bonus (€1.99 – €7.82).

Stimuli. Participants priced monetary gambles of the form "p chance of winning x, otherwise nothing." For the negative condition, these gambles were constructed as follows: 150 random payoffs were drawn from a uniform distribution with the range 1.01–2500. The probabilities for each payoff were set so that they were inversely related to the payoff x ($p = 1 - x/2500$). We jittered payoffs and probabilities by adding normally distributed noise with a standard deviation of 0.1 to both the logit transformation of the probabilities and the logit transformation of normalized payoffs. We then transformed those perturbed values back to the scales used in the experiment. For the positive condition, we used the same gambles as in the negative condition but reversed the order of the probabilities. For the uncorrelated condition, we randomly linked probabilities and payoffs. This approach maintained the marginal distribution of payoffs and probabilities across conditions (as marginal distributions of payoffs and probabilities may influence choice; see Stewart et al., 2006).

In addition to these 150 condition-dependent gambles, participants also priced 22 gambles that were common to each of the three conditions, yielding 172 gamble stimuli per condition. Specifically, we included 10 less extreme gambles in the center of the payoff–probability distribution space (intermediate payoffs and probabilities). These gambles fit equally well within each condition and were not linked to either high or low (i.e., extreme) payoffs. They were therefore used to study condition-dependent differences beyond gambles being surprising, expected, or reference gambles (these gambles are not depicted in Figure 1). and 12 test gambles at the margins of the payoff–probability distribution space. These consisted of 3 high payoff/high probability ($$$,%%%), 3 high payoff/low probability ($$$,%), 3 low payoff/high probability ($,%%%), and 3 low payoff/low probability ($,%) gambles; see triangles in Figure 1). Payoffs were random draws between E$1–E$500 (low) and E$2000–E$2500 (high). Probabilities were random draws between 0.01–.2 (low) and .8–1.0 (high). Payoffs and probabilities were factorially combined to obtain the four payoff/probability combinations. The test gambles were

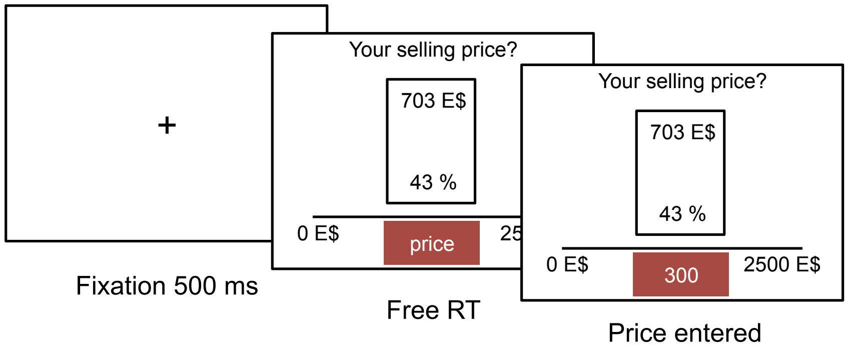

Figure 2. Task. In Experiment 1, participants saw a gamble and priced it in their own time. The price was set by moving an arrow along a rating scale and confirming the price by clicking on it.

interspersed after 100 environment-only trials (i.e., in the last third of the trials).

Procedure. Participants indicated their willingness to sell (WTS) for one gamble at a time (see Figure 2), taking self-paced breaks between the five blocks. The task was presented in the form of a game show called "Keep or sell?" ("Behalten oder Verkaufen?"). To motivate participants to indicate their true valuations of the gambles, we implemented a Becker-DeGroot-Marschak auction (Becker et al., 1964) as follows: Participants entered a price at which they would be willing to sell each gamble by moving the mouse along a rating scale (E$0–2500) and confirming the value with a click. To incentivize the task, the experimenter informed participants that 10 gambles would be randomly selected at the end of the experiment. For those 10 gambles, the experimenter then offered a randomly generated buying price between 0 and the absolute payoff in that gamble. If the experimenter's buying price exceeded the participant's selling price, participants sold the gamble and earned the buying price. If the participant's selling price exceeded the experimenter's buying price, the gamble was played out (e.g., 50% chance of E$380). The dominant strategy in this task is to price a gamble based on its subjective value: Setting higher prices can prevent participants from selling unattractive gambles; setting lower prices can lead to them selling attractive gambles under value. In other words, the prices should approximate participants' certainty equivalents for the gambles. Experiments were coded in PsychoPy (Peirce, 2007).

Statistical Analyses

We used Bayesian estimation techniques (Kruschke, 2015). Specifically, we applied Bayesian generalized linear mixed models using Stan in R for regression analyses with the rstanarm package (*RStanArm*, Version 2.9.0–4). All regression models used trial-level data and participant as a grouping factor. We ran three chains using Markov Chain Monte Carlo sampling to draw from posterior distributions of parameters. Depending on model complexity, we ran 10,000–30,000 samples per chain (to ensure an effective sample size of >10,000 for each regressor) and set a burn-in of 500 samples. We investigated (convergence of) our posteriors through visual inspection and the Gelman–Rubin statistic (Gelman & Rubin, 1992). In general, we report the mean of the posterior distribution of the parameter or statistic of interest and two-sided 95% equal tail credible intervals (CI) around each value. In all analyses, we compared how condition-dependent expectations (expected, surprising, reference) modulated the behavioral measure of interest (deviations from prices, deviations from typical RTs, pupil dilation). We modeled all gamble types simultaneously and used the "$,%" gambles as baseline gambles. For the "expectation" variable, we used "expected" as the baseline. As the RT data were slightly right-skewed, we normalized RTs using a log-transformation of the data before running our analyses. For better interpretability, we report and plot parameters and credible intervals from regression models using untransformed data. Qualitatively, the conclusions were identical for log-transformed and untransformed data (see Open Science Framework).

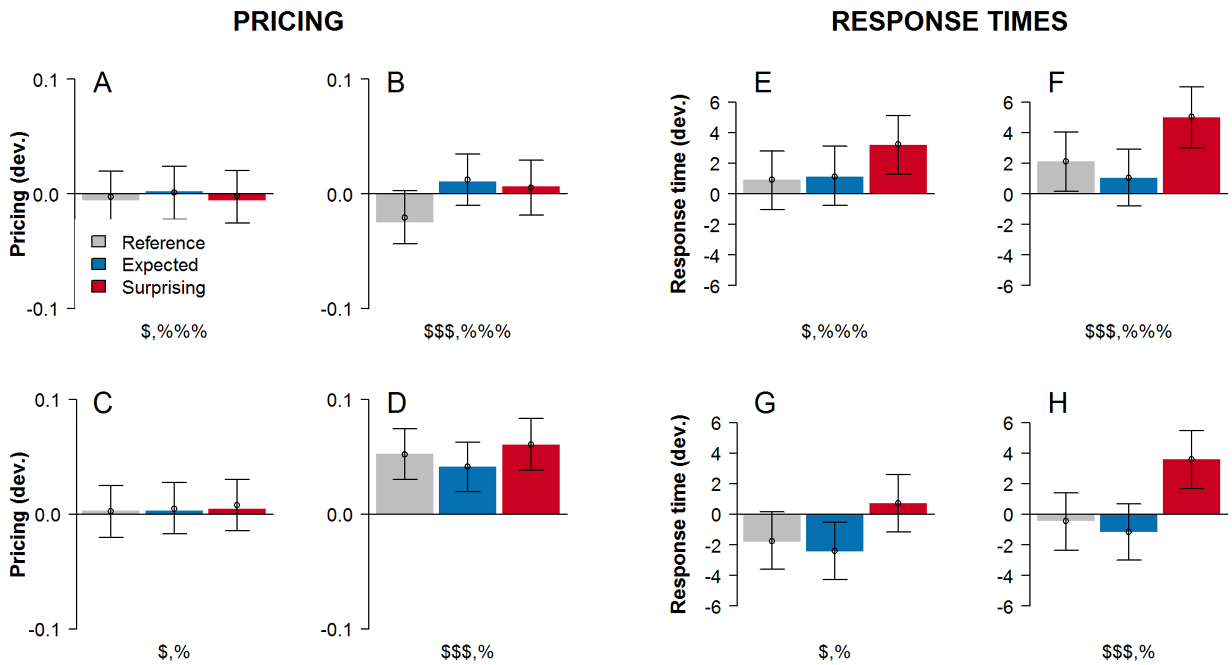

Figure 3. Behavioral results for the test gambles interspersed in Experiment 1. (A–D) Pricing. Prices expressed as normalized deviations from the gambles' expected values ($\frac{price}{2500} - \frac{EV}{2500}$). We found little to no difference in how well prices were adjusted to the expected values of the gambles. (E–H) Response times. Participants slowed down when options were surprising. RTs expressed as deviations from individual median RTs across trials ($RT_{trial} - RT_{ind.\,Md}$). Black dots and error bars represent the mean and the 95% credible interval of the posterior predictive distribution. Reference = uncorrelated condition in all panels; expected = negative condition in panels A, D, E, H/positive condition in panels B, C, F, G; surprising = negative condition in panels B, C, F, G/positive condition in panels A, D, E, H.

Results

We excluded trials in which participants indicated prices that exceeded the payoff offered in the gamble by more than E$100, as this suggests lack of attention to the task. We also removed trials in which RTs deviated by $+/-3SD$ from an individual's median, assuming that these RTs are unlikely to reflect cognitive processing in a specific trial. In total, we removed 4.6% (718/15,480) of the trials across all participants. We analyzed our data with and without these excluded trials and obtained qualitatively very similar results. In the following report, we note when these exclusions led to qualitatively different results.

Prices. Across all gambles, including the environment gambles, prices were strongly related to the gambles' expected values (payoff × probability interaction, $b = 0.90$, CI = $[0.87, 0.92]$, with the uncorrelated condition as baseline). Prices in the negative condition were slightly more adjusted to the expected value than were prices in the other two conditions (payoff × probability × negative condition, $b = .05$, CI = $[.01, .07]$). This could be an artifact of the narrower distribution of expected values in the negative condition than in the other two conditions (which may result in similar price estimates, and less error, throughout). Indeed, there were no differences between conditions when we used normalized deviations from expected values ($\frac{price}{2500} - \frac{EV}{2500}$) as a dependent variable and condition as predictor (rather than modeling prices based on the payoff × probability interaction; all condition-dependent CIs included 0). There were also no condition-dependent differences for test gambles with intermediate payoff–probability combinations (all CIs included 0).

We hypothesized that participants would provide prices closer to the gambles' expected values when risk–reward combinations were expected than when they were surprising. When gambles were surprising, we hypothesized a deviation in the direction of the probability expected from the structure of the environment. However, as Figure 3 shows, there were no reliable pricing differences as a function of whether the gamble was surprising vs. expected ($b = 0.001$, CI = $[-0.010, 0.013]$, surprising vs. expected across gamble types); or whether it was a reference gamble vs. expected ($b = 0.006$, CI = $[-0.009, 0.021]$). There was a small (but noncredible) difference in pricing for gambles that were surprising, but only when gambles offered high payoffs. Consistent with our prediction, prices for high-probability gambles with a surprisingly high payoff ($$$,%%%)

decreased ($b = -0.009$, CI $= [-0.047, 0.029]$, surprising vs. expected). At the same time, prices for low-probability gambles with a surprising high payoff ($$$,%) increased ($b = 0.014$, CI $= [-0.028, 0.057]$, surprising vs. expected).[4] It is plausible that prices were slightly adjusted to what could be expected (a low probability % in $$$,%%% gambles and a high probability %%% in $$$,% gambles).

Response times. The mean RT across gamble types was 15.3 s per trial in the uncorrelated condition (CI $= [13.52, 17.13]$) and credibly lower in the negative ($b = -3.16$, CI $= [-5.75, -0.53]$) and positive ($b = -2.92$, CI $= [-5.51, -0.38]$) conditions.[5] In addition, participants in the uncorrelated and negative conditions responded slightly faster as trials proceeded ($b_{uncor} = -0.013$, CI $= [-0.017, -0.009]$). This pattern was slightly more pronounced in the negative condition ($b_{neg} = -0.008$, CI $= [-0.013, -0.002]$). The effect of trial number on RT was not credible in the positive condition ($b_{pos} = 0.007$, CI $= [-0.002, 0.015]$, with the uncorrelated condition as baseline).[6] There were no condition-dependent RT differences for test gambles with intermediate payoff–probability combinations (all CIs included 0).

How did RTs differ when gambles were surprising? To address this question, we computed individual median RTs across all trials and computed deviations from these median RTs for each trial. These could then be used as an indicator of condition and processing differences dependent on specific gamble types ($RT_{trial} - RT_{ind.\,Md}$). As Panels E–H in Figure 3 show, participants slowed down in response to surprising versus expected

[4] These gambles were also overpriced in general (main effect), possibly a result of the high payoff > E$2000 and insufficient adjustment for low probabilities. Experiment 2 yielded a similar result, as reported later.

[5] When all trials were included in the analysis, participants' average RT decreased to 10.8 s in the uncorrelated condition (CI $= [8.54, 11.57]$) and was again slightly faster in the correlated conditions, $b_{neg} = -1.18$, CI $= [-4.44, -0.008]$, $b_{pos} = -1.94$, CI $= [-5.15, -0.85]$.

[6] The differential effect of the positive condition over the baseline condition (uncorrelated) was $b_{pos} = 0.020$, CI $= [0.015, 0.026]$. When this interval was added to the interval of the uncorrelated condition, the net effect of trial on RTs was no longer credible.

gambles. On average, they spent 3.5 s longer on surprising gambles than on expected gambles ($b = 3.48$, CI = $[2.53, 4.42]$, modeled across gamble types). RTs to reference gambles were statistically indistinguishable from those to expected gambles.

When RTs were broken down by specific payoff–probability combinations, participants were 2.5 s quicker when responding to the least attractive gambles in all conditions, irrespective of whether the gambles were surprising or not ($,%, $b = -2.40$, CI = $[-3.83, -0.97]$). Conversely, all participants took more time responding to the most attractive gambles ($$$,%%%, $b = 3.45$, CI = $[1.57, 5.34]$). Responses were further slowed down by about another second when gambles were also surprising (see Figure 3, panel F), but this effect was not credible ($b = .78$, CI = $[-1.91, 3.49]$). All results were qualitatively the same in a model using normalized, log-transformed RTs.

The most attractive gambles in the negative condition were surprising. Put differently, participants in the *negative* condition took much longer responding to the most attractive gambles ($$$,%%%) than to less attractive, expected gambles ($$$,%). They also took longer responding to these gambles than in the positive condition, where these attractive gambles ($$$,%%%) were expected. A similar effect was observed for participants in the positive condition responding to high-payoff gambles with an unexpectedly low probability ($$$,%, see Figure 3, panel H). [7]

Conclusion

Experiment 1 suggests that people form expectations about risk–reward structures that reflect the experienced environments. This was evident in participants' responses to options that did not match these expectations. Specifically, they slowed down in response to surprising options. At the same time, their adjustments of prices to the gambles'

[7] These results are consistent with a model in which we regressed RTs onto expected values, conditions and their interaction: RTs generally increased as gambles offered higher expected values ($b_{unc} = 1.73$, CI = $[0.84, 2.61]$). However this was much more pronounced in the negative condition ($b_{neg} = 3.40$, CI = $[1.35, 5.40]$), and not credible in the positive condition ($b_{pos} = -0.58$, CI = $[-1.71, 0.54]$).

expected values were the same for all gamble types (surprising, expected, and reference). It is plausible that participants slowed down when options were surprising to achieve the same level of pricing "accuracy" as when options were expected. In addition to the overall RT effects, the data showed that participants' processing of the options was highly dependent on the attractiveness of the gambles: Differences in RTs were more pronounced for surprisingly good ($$$,%%%) than for surprisingly bad gambles ($,%)—to which participants responded faster in general. Notably, these surprise effects emerged irrespective of any feedback on whether or not an outcome was obtained.

Experiment 2: Pupillometric and Behavioral Responses to Surprising Risk–Reward Combinations

In this experiment, we sought to replicate the findings of Experiment 1; in addition, we considered pupil dilation as a physiological measure of surprise. We hypothesized that pupil dilation would increase when participants were presented with gambles offering a surprising combination of risk and reward.

Methods

We adapted the methodology from Experiment 1 to include an eye-tracking component that allowed us to measure pupillary responses to surprising options. We outline key differences below.

Participants. Ninety-three (55 female) participants (age $M = 25.6$ years, $SD = 3.7$ years) from the participant pool at the Max Planck Institute for Human Development, Berlin, completed the experiment (duration 75 min). All participants were paid a fixed rate of €12 plus a bonus based on their performance (€3.53–€11.67).[8]

[8] After the pricing task reported here, participants completed a risky choice task. The respective data and analyses are reported elsewhere (Leuker et al., 2017b). The reported duration includes the additional choice tasks.

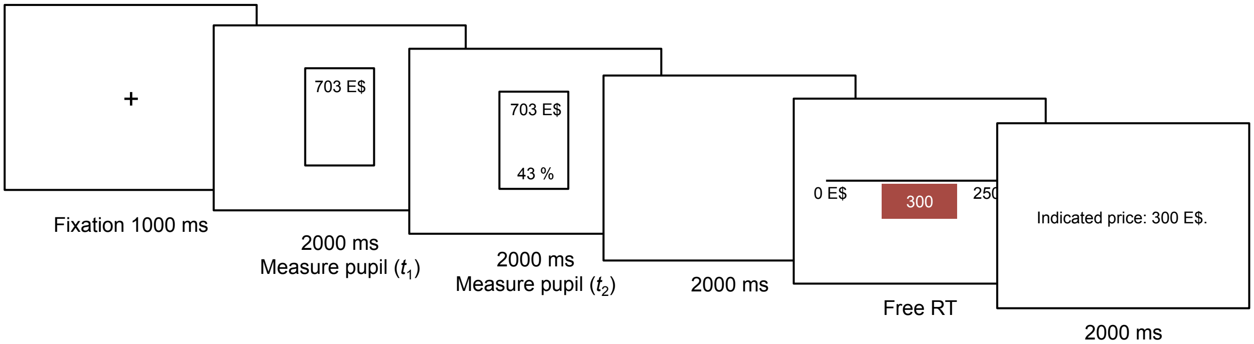

Figure 4. In Experiment 2, participants priced gambles after seeing information on the payoff and probability of each gamble in a fixed sequence. After 4 s in total, the gamble disappeared from the screen. We added a blank screen to achieve sufficient spacing between the critical stimuli (i.e., a gamble's risk–reward relationship) and the participant's manual response. Subsequently, participants had as much time as they wanted to indicate a price for the gamble on a rating scale. We measured pupil dilation as a response to a gamble's payoff alone (t_1) and as a response to its payoff and probability combined (t_2).

Stimuli. We reduced the number of stimuli to 90 condition-dependent gambles and increased the number of gambles common to each of the three conditions. Again, we also added six gambles with intermediate payoff–probability combinations. As these gambles fit equally well within each condition and were not linked to either high or low (i.e., extreme) payoffs, they were used to study condition-dependent differences beyond gambles being surprising, expected, or reference gambles (they are not depicted in Figure 1). As test gambles, we created six gambles for each combination of risks and rewards, or payoffs and probabilities. Overall, these procedures resulted in 120 gambles. As the triangles in Figure 1 show, only a subset of the test gambles was surprising. In Experiment 2, the surprise/environment gamble ratio was (12/120 = .10). Half of the test gambles will always blend in with the condition's environment gambles. The test gambles were interspersed after 40 environment-only trials.

Procedure

Experiment 2 differed from Experiment 1 in that each gamble's payoff and probability appeared sequentially: After a fixation cross, the payoff appeared for 2 s, followed by the probability for another 2 s. After a blank screen (2 s), the screen automatically switched to a rating scale. Participants entered the prices they were willing to sell the gamble for by moving the mouse along this rating scale (E$0–E$2500) and clicking on the value to confirm (free RT). This experimental paradigm can help to disentangle reactions to payoff only, reactions to payoff and probability combined, and manual responses (Figure 4). To control for the pupillary light reflex, we matched the gambles' luminance to the background of the screen by defining stimuli colors in the Derrington, Krauskopf, and Lennie color space (see Derrington et al., 1984; MacLeod & Boynton, 1979). Gambles were presented in orange on a gray background with the same luminance (for clarity's sake, these colors are not shown in the figure).

Eye Tracking

We collected binocular eye tracking data with an EyeTribe tracker, sampled at 60 Hz. The experiment was implemented in PsychoPy 1.83.01 (Peirce, 2007) and the eye-tracking interface PyTribe (Dalmaijer et al., 2014). Before the task, each participant's eye movements were calibrated using the EyeTribe UI with a 9-point grid (< 0.7 degrees of visual angle). Participants were seated approximately 60 cm from the screen with their chin on a chinrest affixed to the table, in a room with negligible ambient light. We obtained pupil size from the left and the right eye (arbitrary units, measured at every Hz).

Eye-Tracking Analyses

Pupillary data were preprocessed as follows. We used EyeTribe's default settings to detect fixations and removed saccade data, because pupillary responses during (and even before) saccades differ systematically from those during fixations (Mathôt et al., 2016).

Trials were discarded when pupil size deviated more than 3 SDs from a participant's median pupil size and when it was outside plausible values (range [10,40] in arbitrary units given by Eyetribe). This procedure removed blinks (rows with values [0,0]) and measurement error. We smoothed the data using a lowess filter, and we averaged the pupil size of the left and the right eye. We removed trials with fewer than 15 samples, which would indicate extremely poor eye tracking (one would expect 240 samples, minus a few blinks, per 2 s period). This applied to a small proportion of all trials: .07 of all trials at t_1 (when reward was shown); .05 of all trials at t_2 (when reward and probability were shown).

To facilitate comparisons across participants and consistent with the literature, we analyzed pupillary signals aligned to a baseline pupil size. As a baseline signal, we used the offset of stimulus presentation (median value in the first 100 ms of a trial).[9] We did so by subtracting the signal at each time point from the baseline signal and then dividing by the baseline signal, resulting in a percentage change relative to the stimulus onset.

Pupillary responses to psychologically relevant stimuli are thought to occur after approximately 1000 ms (and are therefore conceptually different from the pupillary light reflex that occurs after milliseconds, see Gagl et al., 2011; Van Steenbergen & Band, 2013). As in previous research using pupillary responses in the context of choice (Cavanagh et al., 2014), we therefore set an a priori region of interest from 1000 to 2000 ms poststimulus at t_1 (payoff visible on screen) and t_2 (payoff and probability visible together on screen) for our statistical analyses (see Figure 1). We obtained the median percentage change in pupil dilation within this a priori region of interest. The median percentage change is a conservative estimate that may even underestimate the true effect but is more robust to measurement error than is peak dilation. We compared the results of this analysis with results using the mean dilation (e.g., as in Mathôt et al., 2016) and peak dilation (e.g., as

[9] We used this baseline to obtain a similar measure of pupil changes for both t_1 and t_2. Often the fixation cross time is used as a baseline; in our paradigm, however, the fixation cross only preceded stimulus appearance at t_1; see Figure 4.

in Fiedler & Glöckner, 2012), which were qualitatively very similar (see Supplementary Materials for results using the other two indicators). An additional advantage of having a single value for the percentage pupil change per trial is that this analysis automatically controls for multiple testing (due to an otherwise vast number of pupil samples recorded at 60 Hz). We would like to stress that our focus was not on the time course of pupil dilation because we introduced a fixed lag between payoffs, probabilities, and participants' ability to respond (however, for completeness, we plot the timecourse of pupillary responses at t_1 and t_2 in the Supplementary Materials). While some research has studied pupil dilation shortly before a decision is made (Fiedler & Glöckner, 2012, finding that pupil dilation increases as the participant is deciding), in our setup we cannot determine a unique time point at which participants made their choice: They could have reached a decision prior to being able to enter it on the rating scale.

Results

As in Experiment 1, we excluded trials in which participants indicated prices that exceeded the payoff offered in the gamble by more than E$100, as this suggests lack of attention to the task. Moreover, we removed trials in which RTs deviated by $+/-3SD$ from an individual's median, assuming that these RTs are unlikely to reflect cognitive processing on a specific trial. Overall, these standards resulted in the removal of 9.2% (1,027/11,160) of trials across all participants. We analyzed our data with and without these excluded trials and obtained qualitatively very similar results. In the following report, we note when these exclusions led to qualitatively different results.

Pricing. For all participants, prices were strongly related to the expected values of the gambles (credible payoff × probability interaction, $b_{uncor} = 0.86$, CI = $[0.83, 0.89]$). This relationship was slightly more pronounced in the negative condition ($b_{neg} = 0.055$, CI = $[0.010, 0.099]$) and slightly less pronounced in the positive condition ($b_{pos} = -0.051$, CI = $[-0.096, -0.006]$) than in the uncorrelated condition. These differences could be an artifact of the different marginal distributions of expected values across the conditions.[10] When we controlled for expected value, prices deviated positively from the expected value in all three conditions ($b_{unc} = 0.030$, CI = $[0.004, 0.057]$, $b_{neg} = -0.013$, CI = $[-.041, .024]$, $b_{pos} = 0.008$, CI = $[-0.030, 0.048]$, but only credibly in the uncorrelated condition).[11] Prices did not deviate in a particular direction for gambles with intermediate payoff–probability combinations (gambles common across conditions but no test gambles; all CIs included 0).

Across all gamble types, the deviation of prices from expected value differed somewhat between expected and surprising options ($b = 0.012$, CI = $[0.000, 0.023]$). As Figure 5 (panels A–D) shows, the difference between expected and surprising options was driven by high-payoff gambles (panels B, D). Specifically, participants indicated lower prices when the most attractive gambles ($$$,%%%) were surprising (i.e., in the negative condition) than when they were expected (i.e., in the positive condition) ($b = -0.037$, CI = $[-0.069, -0.005]$). Thus, when the most attractive gambles were surprising, prices were adjusted to what could be expected (a low probability). Conversely, participants indicated higher prices when high payoff/low probability gambles ($$$,%) were surprising than when they could be expected ($b = 0.086$, CI = $[0.018, 0.156]$; all models using $% gambles as baseline). Again, prices in the surprising condition were adjusted to what these participants might have expected (a high probability). Taken together, these results are consistent with the probabilities that participants could expect from high payoffs: the negative condition prompts the expectation that a high payoff will be accompanied by a low probability; the positive condition prompts the expectation that a high payoff will be accompanied by a high probability. However, prices did not seem to be adjusted to

[10] When all trials were included in the analysis, prices were (as expected) less well adjusted to gambles' expected values, as indicated by a weaker but credible payoff × probability interaction, $b_{unc} = 0.70$, CI = $[0.66, 0.71]$. Again, there were no differences between conditions (all CIs contained 0).

[11] When all trials were included in the analysis, there was a positive deviation from expected values, $b_{unc} = 0.051$, CI = $[0.023, 0.060]$, slightly less so in the negative condition $b_{neg} = -0.023$, CI = $[-0.063, -0.009]$, $b_{pos} = -0.005$, CI = $[-0.044, 0.008]$.

participants' expectations in their pricing of low-payoff gambles.[12]

[12] When all trials were included in the analysis, the effect of surprise on prices for the attractive gambles ($$$,%%%) was not credible, but in the same direction ($b = -0.017$, CI = $[-0.061, 0.027]$). The effect on the high-payoff, low-probability gambles ($$$,%) was qualitatively identical ($b = 0.10$, CI = $[0.018, 0.020]$).

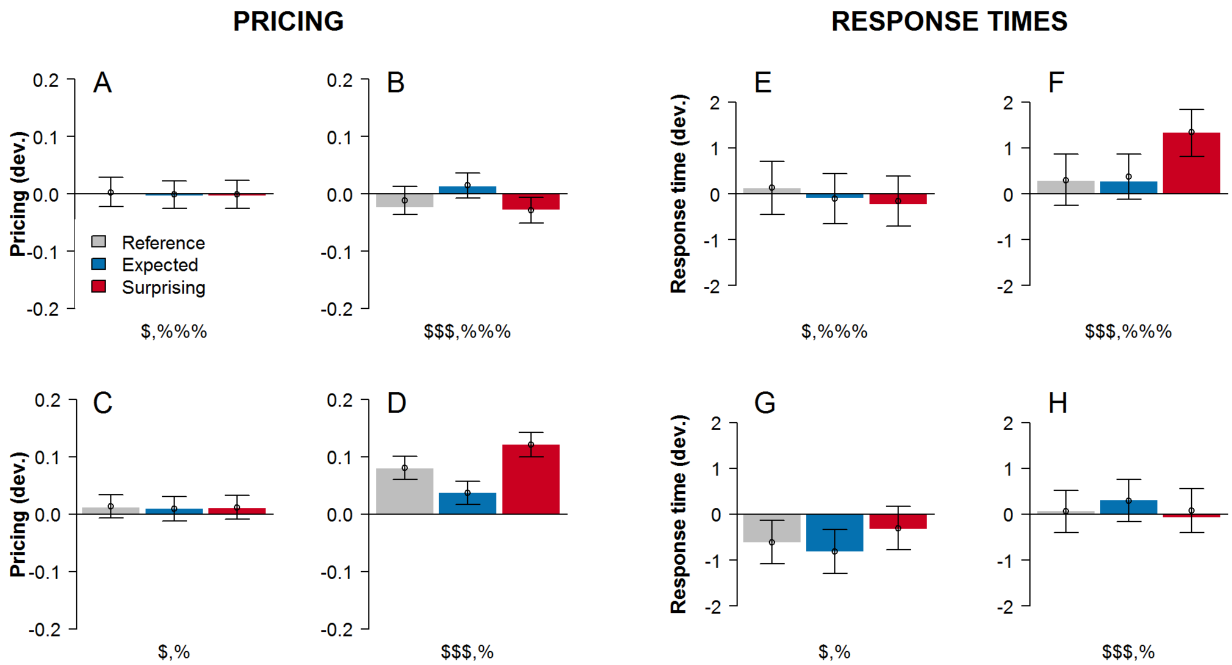

Figure 5. Behavioral results for the test gambles interspersed in Experiment 2. (A–D) Pricing. Prices expressed as normalized deviations from the gambles' expected values ($\frac{price}{2500} - \frac{EV}{2500}$). For low payoffs, we found little to no difference in how well prices were adjusted to the expected values of the gambles. For high payoffs, prices of surprising options differed from prices of expected options. (E–H) Response times. Participants slowed down when options were surprising. This effect was driven by the surprisingly good options (panel F). RTs expressed as deviations from individual median RTs across trials ($RT_{trial} - RT_{ind.\,Md}$). Black dots and error bars represent the mean and the 95% credible interval of the posterior predictive distribution. Reference = uncorrelated condition in all panels; expected = negative condition in panels A, D, E, H/positive condition in panels B, C, F, G; surprising = negative condition in panels B, C, F, G/positive condition in panels A, D, E, H.

Response times. Participants took 4.67 s (CI $= [4.06, 5.28]$, RT model with the uncorrelated condition as baseline) on average per trial from seeing the empty screen (2 s enforced) to entering a response on the rating scale. There were no credible differences in these average RTs across conditions (all CIs for condition-dependent effects contained 0). Similar to Experiment 1, participants' response speeds increased as trials proceeded ($b = -0.008$, CI = $[-0.11, -0.006]$).[13] RTs did not differ across conditions for test gambles with intermediate payoff–probability combinations (all CIs included 0).

The range of RTs was smaller in Experiment 2 than in Experiment 1 due to the experimental design, in which a price could not be entered before seeing payoff and probability information for 2 s each. However, as in Experiment 1, our focus was on RT differences for particular gamble types. Again, RTs varied depending on the payoff–probability combination: Consistent with Experiment 1, participants across all conditions responded faster to unattractive gambles ($,%, $b = -0.55$, CI = $[-1.05, -0.04]$) and spent more time evaluating the most attractive gambles ($$$,%%%, $b = 0.65$, CI = $[0.19, 1.11]$, main effects).

Moreover, RTs again depended on the gamble type: In Experiment 2, participants spent 0.28 s longer on surprising gambles than on expected gambles (CI = $[0.031, 0.52]$). As in Experiment 1, RTs for reference gambles were indistinguishable from those for expected gambles. As Figure 5 shows, the effects of surprise differed slightly depending on the gambles' payoff–probability combinations. In particular, surprise elevated the RT differences for the most attractive gambles, but this effect was not credible ($$$,%%%, $b = 0.40$, CI = $[-0.28, 0.63]$, expectation × gamble type). These differences in RTs were qualitatively identical when data from all participants was entered in the analysis.[14] [15]

---

[13] These "practice effects" disappeared when all participants were included in the analysis. Likewise, they did not emerge in the positive and negative conditions when we used normalized, log-transformed RTs.

[14] All results were qualitatively the same in a model using normalized, log-transformed RTs.

[15] As in Experiment 1, these results are consistent with a model in which we regressed RTs onto expected values, conditions and their interaction: RTs generally increased as gambles offered higher expected values ($b_{unc} = 0.34$, CI = $[0.00, 0.69]$). Again, the link was much more pronounced in the negative condition ($b_{neg} = 1.41$, CI = $[0.81, 2.01]$), and (also as before) not credible in the positive condition ($b_{pos} = 0.20$, CI

In sum, both the payoff a gamble offers and whether or not it is surprising influence the pricing and processing of gambles. In Experiment 2, participants spent more time evaluating attractive gambles (%%%,$$$) when those gambles were surprising (i.e., in the negative condition). However, the differences in RTs were much smaller than in Experiment 1. This difference is likely an artifact of the experimental design, in which a high proportion of the overall RT was fixed (payoff presented for 2s, payoff + probability for 2 s, blank screen for 2 s = at least 6 s to process the gamble before being able to enter a price).

Pupil dilation. We analyzed pupil responses at two time points. Pupillary responses to payoffs only (t_1) were used to test the influence of payoff magnitude: If participants responded to surprising options (defined as payoff–probability combinations), pupil size should not vary between high or low payoffs at t_1. This was indeed the case across all conditions (all CIs contained 0).[16] How did pupil dilation change as a function of whether or not a gamble was surprising? Figure 6 shows mean changes in pupil dilation after both payoff and probability information was presented (t_2). Panels A–D suggest that pupil size was not associated with gambles being either surprising or expected ($b = -0.14$, CI = $[-0.92, 0.63]$, main effect of surprising vs. expected across gamble types), nor with differences between expected and reference gambles ($b = 0.06$, CI = $[-1.01, 1.12]$, main effect of reference vs. expected across gamble types). Instead, pupil size was affected by surprise and how attractive a gamble was. Specifically, pupil size decreased in response to surprisingly bad gambles (Panel C: $,%, $b = -1.75$, CI = $[-3.38, -0.11]$) and increased in response to surprisingly good gambles (Panel D: $$$,%%%, $b = 2.84$, CI = $[0.75, 4.96]$). When these gambles were expected or when they were reference gambles (i.e., in the uncorrelated condition), pupil responses were statistically indistinguishable from 0 (all CIs

---

= $[-0.23, 0.65]$).

[16] $b_{unc} = -0.07$, CI = $[-0.91, 0.76]$, using the uncorrelated condition as baseline; $b_{neg} = -0.40$, CI = $[-1.59, 0.71]$, $b_{pos} = -0.33$, CI = $[-1.52, 0.88]$, all comparisons $$$ > $. These results are plotted in the Supplementary Materials. The results were qualitatively identical with and without excluded trials.

contained 0). Panels A and D show that the change in pupil dilation was statistically indistinguishable from a 0% change for gamble types involving some tradeoff between payoffs and probabilities ($$$,% or $,%%%) across all three conditions—that is, irrespective of whether these gambles were surprising or not.

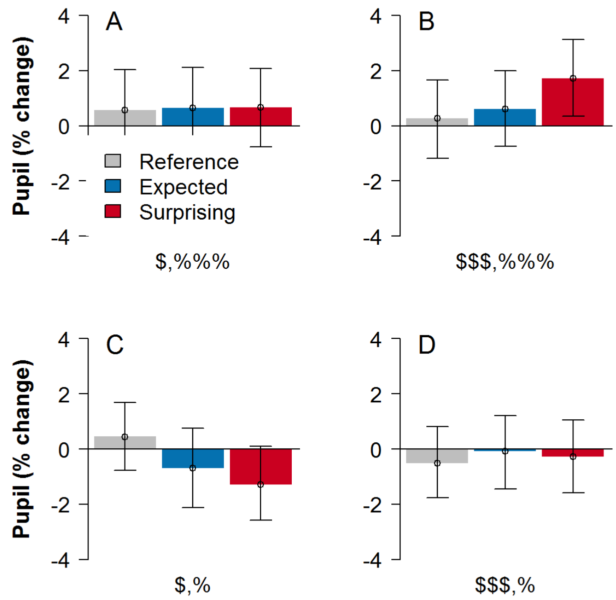

Figure 6. Pupil dilation after seeing gambles (both payoff and probability information on the screen). Pupil dilation computed as the median percentage change from 1000–2000 ms after the stimulus. Black dots and error bars represent the mean and the 95% posterior predictive distribution.

Conclusion

Experiment 2 further corroborates that people seem to build up expectations about the options appearing in a particular risk–reward environment. As in Experiment 1, this was evident in longer RTs for surprising gambles. Moreover, the data from Experiment 2 suggest differences in pricing—but only when gambles offered a high payoff. Specifically, when the probability of such gambles was low but participants expected a high probability given the risk–reward environment, participants indicated prices that were drawn towards that expectation (and were slightly higher). Conversely, when the probability was high but participants expected a low probability, prices were drawn towards that expectation (and were slightly lower). Lastly, surprisingly good options were associated with a reliable increase in pupil size ($$$,%%%). However, pupil size was not associated with surprise per se (across gamble types).

General Discussion

If something sounds too good to be true, it usually is. The only way to identify such *exceptional* options is being sensitive to some of the tradeoffs people typically face in their choice environments. A frequent regularity people encounter is an inverse relationship between risks and rewards (Pleskac & Hertwig, 2014). Using a novel preferential oddball paradigm, we found evidence that (a) people are fairly sensitive to such regularities and (b) they show distinct responses to surprising options that deviate from the regularities. That is, people seem to compare new, incoming options against the risk–reward relationship they learned from their environments. In the preferential oddball paradigm, people learned about risk–reward structures by pricing monetary gambles drawn from negative, positive, or uncorrelated risk–reward environments (also see Leuker et al., 2017a). After some exposure to a particular risk–reward environments, participants encountered surprising options that deviated from learned risk-reward structures. We investigated to what extent these surprising options were linked to three indicators of surprise: prices, RTs, and pupil dilation. Next, we discuss these three indicators in detail. We then relate our findings to previous research on surprise and consider their broader implications for preferential choice.

Do people *evaluate* surprising options differently from nonsurprising options? A comparison across experiments suggests that the time decision makers have available to inspect an option plays an important role. Without timing restrictions (as in Experiment 1, Figure 2), people may have been able to compensate for the uncertainty that surprising options induce simply by looking at or evaluating the options for longer. This may explain the limited differences in pricing we observed when comparing surprising with expected options or reference gambles. In Experiment 2, the time available for inspecting options was limited. Here, prices deviated from gambles' expected values in the direction of participants' expectations: When the probability of surprising gambles was low but participants expected it to be high, prices were drawn towards that expectation (and were slightly higher). Conversely, when the probability was high but participants expected it to be low, prices were drawn towards that expectation (and were slightly lower).[17] However, prices only deviated for options offering high payoffs. While this systematicity speaks against random variability in indicating prices for surprising options, the extent to which choices are affected by surprise may depend on whether or not the duration with which payoffs and probabilities can be inspected is fixed (as in Experiment 2, Figure 4), and the payoffs the surprising options offers.

How do people *process* surprising options? Our data show that people may take more time when evaluating surprising options than when the same options are to be expected. How much longer people take to evaluate a surprising option also depends on whether the option is surprisingly good ($$$,%%%) or surprisingly bad ($,%). This stakes effect was also present in pupil size, which increased in response to surprisingly good options (relative to baseline), but not in response to surprise in general, or in response to high expected value in general ($$$,%%%). The differences between surprisingly good and surprisingly bad gambles are consistent with a "hold your horses" mechanism by which people behave less impulsively in the presence of high–value options (Cavanagh et al., 2014; Frank et al., 2007). That is, more scrutiny is required only when high payoffs are at stake. Here, we have shown that such a mechanism can be weakened when high payoffs are to be expected

[17] We interpret these findings as indicating that participants adjusted probability information to payoffs because payoffs were presented first in Experiment 2, followed by probabilities. This does not rule out the possibility that people may also use the risk–reward relationship in other situations to adjust payoffs to a plausible range.

in an environment, and strengthened when high payoffs are not to be expected.

More generally, our data bring a new perspective to the growing body of research on how the environmental distribution of monetary payoffs and probabilities influences how (otherwise identical) options are evaluated (Birnbaum, 1992; Stewart et al., 2006, 2015; Walasek & Stewart, 2015). Here we show that people go beyond evaluating options from givens. Instead, our results point to a mixture between evaluations from givens and evaluations from environments. For instance, prices for surprising options were adjusted to the options' expected values but sometimes shifted in the direction of environmental expectations. What is more, we observed substantial shifts in the way that surprising options were processed. Such environment-based evaluations cannot be anticipated by prominent theories of choice, which conceptualize risks and rewards as independent attributes that determine the expected utility of an option (von Neumann & Morgenstern, 1944) or its subjective worth (Tversky & Kahneman, 1992).

The fact that people seem to engage in environment-based evaluations (i.e. compare options to the "global" choice environment these options are drawn from) can bring a new dimension to other decision making concepts. Consider the case of dominance. Typically dominance (where one option is superior to all other options on all attributes) or the related concept of stochastic dominance are defined within a given ("local") choice set (Hadar & Russell, 1969). However, there may be a more global aspect to dominance. For example, in the negative risk–reward environment, surprisingly good options *strongly* dominated all other options in the set. Conversely, the surprisingly bad options *were* dominated by all other options in the set. The fact that in our data people showed distinct responses to these value-based "oddballs" (but not to the same options when these could be expected) suggests that they may be sensitive to dominance in these larger, global set of options. Given the important role dominance plays in context effects like the attraction effect (Huber et al., 1982), it is plausible that context effects may not only arise in local choice environments, but also in global choice environments.

In addition, our results bring a novel dimension to studies of surprise. Similar to oddball paradigms using lower level auditory stimuli, we show that feedback is not a prerequisite for the mind to detect surprising stimuli (or deviants) in a sequence of standards. As mentioned before, a crucial difference between oddball paradigms and the *preferential* oddball paradigm we used here is that the latter entails value-based stimuli. Our data suggest that, in contrast to oddball paradigms, the direction of the surprise (i.e., whether an option is surprisingly good or surprisingly bad) matters in preferential choice. Another critical difference is that attention to the stimuli is not required in oddball paradigms (Garrido et al., 2009), but is probably required in the preferential oddball paradigm reported here. By extension, we a "mismatch negativity" may not necessarily emerge for preferential oddballs—a prediction that is directly testable in further research. From this perspective, the results may be more comparable to previous studies on "risk-prediction errors" (Bossaerts, 2010; O'Neill & Schultz, 2013; Preuschoff et al., 2008) or "reward-prediction errors" (Schultz, 2002; Schultz et al., 1997) than oddball paradigms. However, these prediction errors have to date only been elicited in contexts in which feedback is given. Is feedback required for prediction errors to be elicited? This question could be investigated by comparing neurophysiological correlates of surprisingly good versus surprisingly bad options with and without explicit feedback.

A final question pertains to the role of surprise in risk–reward environments. Is there an adaptive aspect to surprise in risk–reward environments? Is its role similar to, or distinct from, other types of surprise? In value-based choice, "reward-prediction errors" are thought to aid reward-based learning. A similar mechanism would be plausible if participants anticipated the surprisingly good option being added to their bonus at the end of the study (without feedback after each choice). Similarly, in oddball paradigms, the mismatch negativity has been described as a marker for error detection that stems from a break in a learned regularity. According to the model adjustment hypothesis, the prediction error leads to a subsequent updating of the [previously learned] model (Garrido et al., 2009).

The model adjustment hypothesis is consistent with the predictive coding framework (Friston & Kiebel, 2009). Adjusting one's model of the world in this way could also be adaptive in risk–reward environments in which the relationships may vary (e.g., a weaker relationship is expected in newly forming markets). However, it remains an open question how much evidence is needed for a model to be adjusted. Our previous experiments suggest that the risk–reward regularity people extract largely reflects the structure of the environment trials (Leuker et al., 2017a). Maintaining a risk–reward rule even after seeing surprising risk–reward combinations would be consistent with a "rule-plus-exception model," according to which people may learn exceptions to a rule instead of updating that rule if they are unable to identify the rule that would account for the exceptions (Nosofsky et al., 1994). For instance, a person "might regard a single-dimension rule as tentatively acceptable as long as it correctly classifies 60% of the incoming exemplars" (p. 56).

In value-based decision making, extracting a rule from the overall environment and forming expectations may serve a very particular function: to help people to identify when options are too good to be true. Forming expectations in this way may work in many nonlaboratory environments in which risks and rewards are inversely related (Pleskac & Hertwig, 2014). In these environments, people know that there is usually "no free lunch," in that the larger rewards they desire occur only rarely (but if they are lucky and get a "free lunch" once, it does not mean that their model of risk–reward environments will change). Ultimately, being sensitive to an environment's risk–reward structures and deviations from those structures can lead to adaptive decisions under risk.

General Conclusion

Three main conclusions can be drawn from the experiments presented. First, people build expectations about the structures of the options in their global choice environments. When presented with options that deviate from these expectations, people slow down to evaluate them. Second, the direction of deviation matters: Surprisingly good options are evaluated with more scrutiny than surprisingly bad options. Third, the expectations people have can directly influence their evaluations of options, particularly when the time available to process those options is fixed. These behavioral signatures may reflect how people deal with surprising options in nonlaboratory environments, especially when those options are "too good to be true": They take some more time to scrutinize them.

References

Becker, G. M., Degroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, *9*(3), 226–232. doi: 10.1002/bs.3830090304

Birnbaum, M. H. (1992). Violations of monotonicity and contextual effects in choice-based certainty equivalents. *Psychological Science*, *3*(5), 310–314. doi: 10.1111/j.1467-9280.1992.tb00679.x

Bossaerts, P. (2010). Risk and risk prediction error signals in anterior insula. *Brain Structure and Function*, *214*(5), 645–653. doi: 10.1007/s00429-010-0253-1

Brooks, D. N. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 169–215). Hillsdale, NJ: Lawrence Erlbaum Associates.

Busemeyer, J. R., & Diederich, A. (2002). Survey of decision field theory. *Mathematical Social Sciences*, *43*(3), 345–370. doi: 10.1016/S0165-4896(02)00016-1

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

Dalmaijer, E. S., Mathôt, S., & Van der Stigchel, S. (2014). PyGaze: An open-source, cross-platform toolbox for minimal-effort programming of eyetracking experiments. *Behavior Research Methods*, *46*, 913–921. doi: 10.3758/s13428-013-0422-2

Derrington, A., Krauskopf, J., & Lennie, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *Journal of Physiology*, *357*, 241–265. doi: 10.1113/jphysiol.1984.sp015499

Dulany, D. E., Carlson, R. A., & Dewey, G. I. (1984). A case of syntactical learning and judgment: How conscious and how abstract? *Journal of Experimental Psychology: General*, *113*(4), 541–555. doi: 10.1037/0096-3445.113.4.541

Fiedler, S., & Glöckner, A. (2012). The dynamics of decision making in risky choice: An eye-tracking analysis. *Frontiers in Psychology*, *3*, 335. doi: 10.3389/fpsyg.2012.00335

Fortin, J. (2017, December 30). Glitch in South Carolina lottery could mean $19.6 million in winnings. *New York Times*. Retrieved from https://www.nytimes.com/2017/12/30/us/lottery-south-carolina.html?smid=pl-share

Frank, M. J., Samanta, J., Moustafa, A. A., & Sherman, S. J. (2007). Hold your horses: Impulsivity, deep brain stimulation, and medication in parkinsonism. *Science*, *318*(5854), 1309–1312. doi: 10.1126/science.1146157

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1211–1221. doi: 10.1098/rstb.2008.0300

Gagl, B., Hawelka, S., & Hutzler, F. (2011). Systematic influence of gaze position on pupil size measurement: Analysis and correction. *Behavior Research Methods*, *43*(4), 1171–1181. doi: 10.3758/s13428-011-0109-5

Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: A Review of underlying mechanisms. *Clinical Neurophysiology*, *120*(3), 453–463. doi: https://doi.org/10.1016/j.clinph.2008.11.029

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*(4), 457–511.

Hadar, J., & Russell, W. R. (1969). Rules for Ordering Uncertain Prospects. *American Economic Review*, *59*(1), 25. doi: 10.1126/science.151.3712.867-a

Harrington, R. (2017, December 31). A computer glitch on Christmas could give mistaken winners in the South Carolina lottery almost $20 million. *Business Insider*. Retrieved from http://www.businessinsider.de/computer-glitch-20-million-dollars-winnings-south-carolina-lottery-2017-12

Herrmann, C. S., & Knight, R. T. (2001). Mechanisms of human attention: Event-related potentials and oscillations. *Neuroscience & Biobehavioral Reviews*, *25*(6), 465–476. doi: 10.1016/S0149-7634(01)00027-6

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

Huettel, S. A., & McCarthy, G. (2004). What is odd in the oddball task? Prefrontal cortex is activated by dynamic changes in response strategy. *Neuropsychologia*, *42*(3), 379–386. doi: 10.1016/j.neuropsychologia.2003.07.009

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–291. doi: 10.2307/1914185

Kruschke, J. (2015). *Doing bayesian data analysis: A tutorial introduction with R.* London, United Kingdom: Academic Press.

Leuker, C., Pleskac, T. J., Pachur, T., & Hertwig, R. (2017a). Exploiting risk–reward structures in decision making under uncertainty (Manuscript submitted for publication). Retrieved from psyarxiv.com/z4knb.

Leuker, C., Pleskac, T. J., Pachur, T., & Hertwig, R. (2017b). How the mind exploits risk–reward structures in decisions under risk. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar (Eds.), *Proceedings of the 39th Annual Conference Cognitive Science Society* (pp. 2543–2548). London, UK.

Love, B. C. (2002). Comparing supervised and unsupervised category learning. *Psychonomic Bulletin & Review*, *9*(4), 829–835. doi: 10.3758/BF03196342

MacLeod, D. I. A., & Boynton, R. M. (1979). Chromaticity diagram showing cone excitation by stimuli of equal luminance. *Journal of the Optical Society of America*, *69*(8), 1183–1186. doi: 10.1364/JOSA.69.001183

Mathôt, S., van der Linden, L., Grainger, J., & Vitu, F. (2016). The pupillary light response reflects eye-movement preparation. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(1), 28–35. doi: dx.doi.org/10.1037/a0038653

Mellers, B. A., Schwartz, A., Ho, K., & Ritov, I. (1997). Decision affect theory: Emotional reactions to the outcomes of risky options. *Psychological Science*, *8*(6), 423–429. doi: 10.1111/j.1467-9280.1997.tb00455.x

Näätänen, R. (2007). The mismatch negativity: Where is the big fish? *Journal of Psychophysiology*, *21*(3–4), 133–137. doi: 10.1027/0269-8803.21.34.133

Nelson, D. G. K. (1984). The effect of intention on what concepts are acquired. *Journal of Verbal Learning and Verbal Behavior*, *23*(6), 734–759. doi: 10.1016/S0022-5371(84)90442-0

Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. , *101*(1), 53–79. doi: 10.1037/0033-295X.101.1.53

O'Neill, M., & Schultz, W. (2013). Risk prediction error coding in orbitofrontal neurons. *The Journal of Neuroscience*, *33*(40), 15810–15814.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker.* New York, NY: Cambridge University Press.

Peirce, J. W. (2007). PsychoPy - Psychophysics software in python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13. doi: 10.1016/j.jneumeth.2006.11.017

Picton, T. W. (1992). The P300 wave of the human event-related potential. *Journal of Clinical Neurophysiology*, *9*(4), 456–479.

Pleskac, T. J., & Hertwig, R. (2014). Ecologically rational choice and the structure of the environment. *Journal of Experimental Psychology: General*, *143*(5), 2000–2019. doi: 10.1037/xge0000013

Preuschoff, K., Quartz, S. R., & Bossaerts, P. (2008). Human insula activation reflects risk prediction errors as well as risk. *Journal of Neuroscience*, *28*(11), 2745–2752. doi: 10.1523/JNEUROSCI.4286-07.2008

Preuschoff, K., 't Hart, B. M., & Einhäuser, W. (2011). Pupil dilation signals surprise: Evidence for noradrenaline's role in decision making. *Frontiers in Neuroscience*, *5*, 115. doi: 10.3389/fnins.2011.00115

RStanArm. (Version 2.9.0–4). Retrieved from cran.r-project.org/web/packages/rstanarm/

Rüsseler, J., & Rösler, F. (2000). Implicit and explicit learning of event sequences: Evidence for distinct coding of perceptual and motor representations. *Acta Psychologica*, *104*(1), 45–67. doi: 10.1016/S0001-6918(99)00053-0

Savage, L. J. (1954). *The foundations of statistics.* New York, NY: John Wiley & Sons.

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron*, *36*(2), 241–263. doi: https://doi.org/10.1016/S0896-6273(02)00967-4

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

South Carolina may owe double in Christmas lottery glitch. (2018, January 23). *Star Tribune.* Retrieved from http://www.startribune.com/south-carolina-may-owe-double-in-christmas-lottery-glitch/470738623/

Squires, N. K., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology*, *38*(4), 387–401. doi: 10.1016/0013-4694(75)90263-1

Stewart, N., Chater, N., & Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology*, *53*(1), 1–26. doi: 10.1016/J.Cogpsych.2005.10.003

Stewart, N., Reimers, S., & Harris, A. J. (2015). On the origin of utility, weighting, and discounting functions: How they get their shapes and how to change their shapes. *Management Science*, *61*(3), 687–705. doi: 10.1287/mnsc.2013.1853

Todd, P. M., Gigerenzer, G., & the ABC Research Group. (2012). *Ecological rationality: Intelligence in the world*. New York, NY: Oxford University Press.

Trueblood, J. S., Brown, S. D., & Heathcote, A. (2014). The multiattribute linear ballistic accumulator model of context effects in multialternative choice. *Psychological Review*, *121*(2), 179–205. doi: 10.1037/a0036137

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323. doi: 10.1007/BF00122574

Usher, M., & McClelland, J. L. (2004). Loss aversion and inhibition in dynamical models of multialternative choice. *Psychological Review*, *111*(3), 757–769. doi: 10.1037/0033-295X.111.3.757

Van Steenbergen, H., & Band, G. (2013). Pupil dilation in the Simon task as a marker of conflict processing. *Frontiers in Human Neuroscience*, *7*, 215. doi: 10.3389/fnhum.2013.00215

von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Walasek, L., & Stewart, N. (2015). How to make loss aversion disappear and reverse: Tests of the decision by sampling origin of loss aversion. *Journal of Experimental Psychology: General*, *144*(1), 7–11. doi: 10.1037/xge0000039

Ward, T. B., & Scott, J. (1987). Analytic and holistic modes of learning family-resemblance concepts. *Memory & Cognition*, *15*(1), 42–54. doi: 10.3758/BF03197711

Wattenmaker, W. D. (1991). Learning modes, feature correlations, and memory-based categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(5), 908–923. doi: 10.1037//0278-7393.17.5.908

Whittlesea, B. W. (1987). Preservation of specific experiences in the representation of general knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(1), 3–17. doi: 10.1037/0278-7393.13.1.3

Supplemental Material for "Too Good to Be True?"

Christina Leuker, Thorsten Pachur, Ralph Hertwig and Timothy J. Pleskac

Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany

January 24, 2018

Contents

1 Pupil dilation (mean vs. peak vs. median)

We report changes in median pupil size in the main text. We obtain qualitatively very similar results when using the mean change (Figure S1), the peak changes (Figure S2) or median changes (Figure S3) in pupil dilation as an indicator. Supplementary figures also show results for gambles with intermediate payoff–probability combinations ($,$,%,%; these were payoffs around 1250E$ and probabilities around 50%). We used these gambles as a different type of reference class as these gambles matched the expectations in all three conditions. Just as the other gamble types, six of these gambles were interspersed after 40 environment–only trials. As expected, there were no credible differences in pupil size for these gambles.

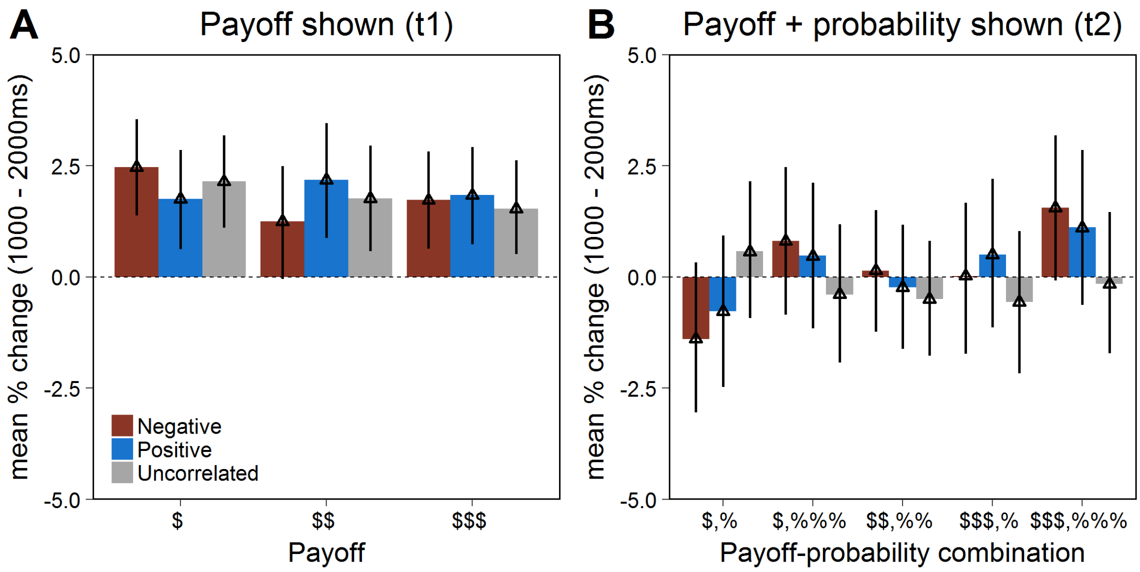

Figure S1: Mean changes in pupil dilation between 1000 and 2000ms after observing (A) payoffs only, (B) or payoff–probability combinations. Black triangles and error bars represent the mean and the 95% credible interval of the posterior predictive distribution.

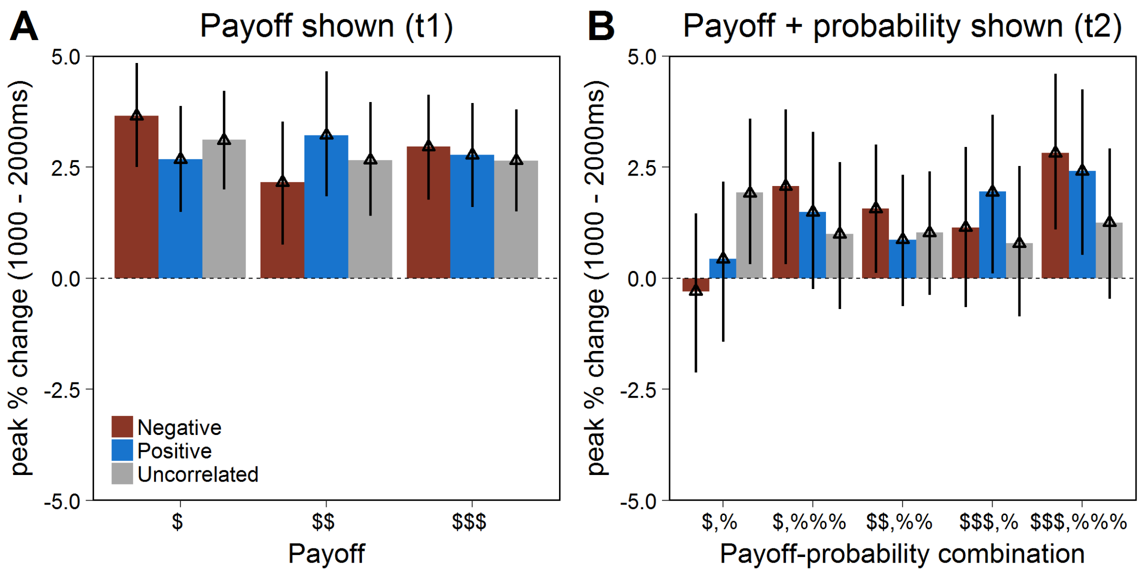

Figure S2: Peak changes in pupil dilation between 1000 and 2000ms after observing (A) payoffs only, (B) or payoff–probability combinations. Black triangles and error bars represent the mean and the 95% credible interval of the posterior predictive distribution.

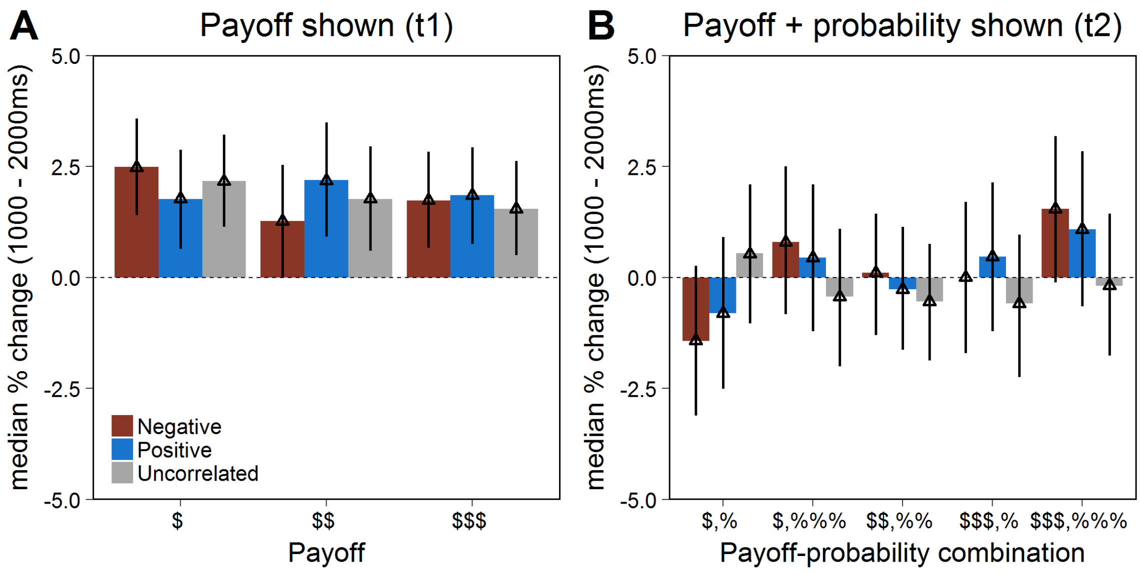

Figure S3: Median changes in pupil dilation between 1000 and 2000ms after observing (A) payoffs only, (B) or payoff–probability combinations. Black triangles and error bars represent the mean and the 95% credible interval of the posterior predictive distribution. Results shown in panel B are reported in the main manuscript.

2 Pupil dilation (temporal dynamics)

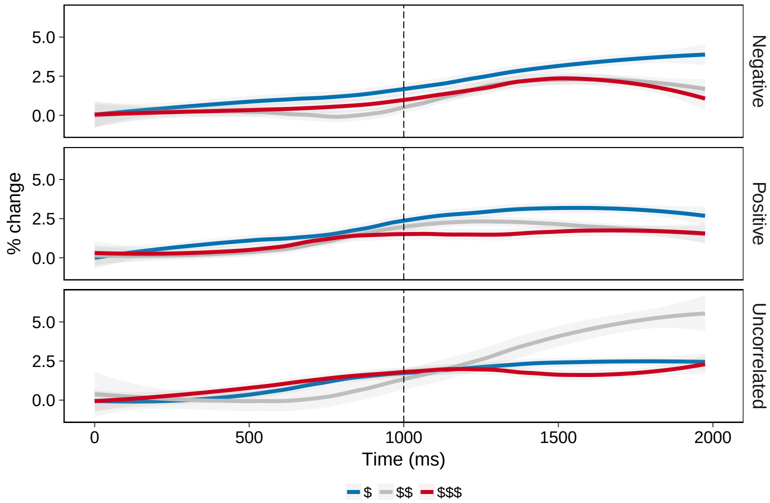

Figure S4: Temporal dynamics of pupil dilation for different payoffs across the three conditions, before any probability was paired with these payoffs. The grey line depicts intermediate payoff levels around 1250E$, paired with intermediate probability levels around 50% (not discussed in main manuscript). Pupil dilation aggregated across participants and trials and smoothed with a loess function.

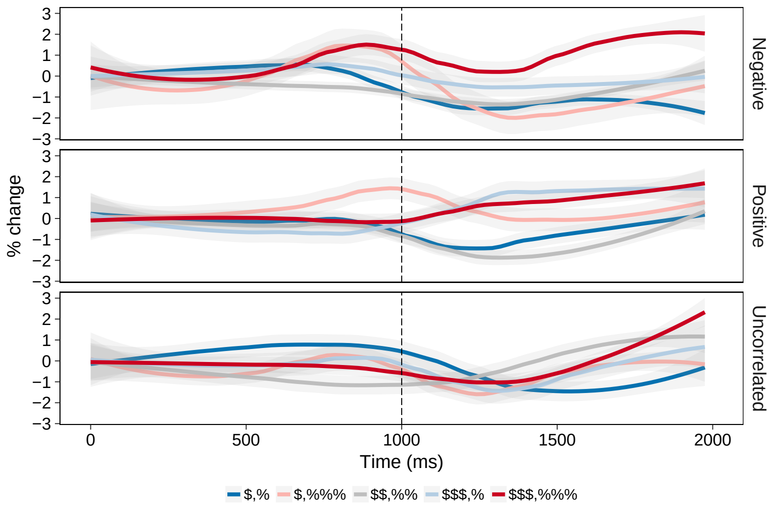

Figure S5: Temporal dynamics of pupil dilation for different payoff–probability combinations across the three conditions. The most attractive gambles in the set are plotted in red, the least attractive gambles are plotted in blue. Pupil dilation aggregated across participants and trials and smoothed with a loess function.

We plotted the change in pupil dilation in response to both only the payoffs (Figure S4) and payoff-probability combinations (Figure S5). Figure S4 shows that, if anything, low payoffs led to a larger % change in pupil dilation than high payoffs (at t1). However, these differences were not statistically meaningful. Figure S5 shows the effects of payoff-probability combinations (at t2). We expected that pupils would be larger in response to surprising gambles. However, across all three conditions, the most attractive gambles ($$$, %%%) are associated with the largest changes in pupil size. The change in pupil size in response the most attractive gambles ($$$, %%%) gambles was credible when these gambles were surprising (in the negative condition).

3 Pupil dilation (exploratory analyses)

3.1 Is pupil dilation linked to expected value?

What is the effect of expected value on pupil dilation across conditions, using all gambles (including environment gambles)? Note that due to the design of the experiment, the (marginal) distribution of expected values differs across the three conditions. Specifically, in the negative condition, expected values never exceeded E$600, except for the six most attractive gambles ($$$, %%%), with expected values close to E$2500. These gambles appeared after 40 environment-only trials. Therefore, the few high EV gambles in the negative condition stood in contrast to all other gambles in the condition with no intermediate EVs (EVs were below E$600 or very high), which might have produced a stronger reaction to high versus low EVs ("salience"). Indeed pupil dilation was linked to EV in the negative condition ($b = .0009$, CI $= [.0003, .0014]$). The marginal distribution of EVs in the positive condition was also bimodal. Again, expected value was linked to pupil dilation ($b = .0006$, CI $=$ $[.0003, .009]$). For both of these correlated conditions, the effect was small. Expected value was not linked to pupil dilation in the uncorrelated condition ($b = .000$, CI $= [-.0004, .0003]$).

3.2 Is pupil dilation linked to response time differences?

Recall that participants slowed down when indicating prices to the most attractive gambles ($$$, %%%) versus least attractive gambles ($, %), especially when these gambles were surprising. The pupillometric data showed differentiable responses to attractive versus unattractive gambles. Are these two processes linked? There was no association between response times and the pupil effect when these gambles were reference gambles ($b_{reference} = 0.21$, CI = $[-.64, 1.07]$), or when these gambles were expected ($b_{expected} = -0.03$, CI = $[-0.82, 0.75]$). When these gambles were surprising, pupillometric responses towards very attractive versus very unattractive gambles are slightly more dissociable in longer trials (that is, when participants slow down to process surprising gambles). Specifically, there was a weak but not credible interaction between taking more time on a given trial, gamble type and surprise ($b = 0.35$, CI = $[-0.14, 0.84]$, model on the trial level with participant as a grouping factor). The effect is slightly larger, but also not credible, when including all participants in the analysis ($b = 0.39$, CI = $[-0.08, 0.87]$).

This analysis is not meaningful for the other payoff–probability combinations ($$$,% and $, %%%, or the reference gambles) — there were little to no (main) effects of these gamble types on reaction times and pupil dilation. That is, there was no association between pupil dilation and reaction times for these gamble types in the first place.