Journal of Behavioral Decision Making, *J. Behav. Dec. Making* (2017)
Published online in Wiley Online Library (wileyonlinelibrary.com) **DOI:** 10.1002/bdm.2007

Inducing Thought Processes: Bringing Process Measures and Cognitive Processes Closer Together

MICHAEL SCHULTE-MECKLENBECK,[1,2]* ANTON KÜHBERGER,[3] BENJAMIN GAGL[4] and FLORIAN HUTZLER[3]

[1]*Department of Business Administration, Consumer Behavior, University of Bern, Bern, Switzerland*
[2]*Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany*
[3]*Centre of Cognitive Neuroscience & Department of Psychology, University of Salzburg, Salzburg, Austria*
[4]*Department of Psychology & Center for Individual Development and Adaptive Education of Children at Risk (IDeA), Goethe University Frankfurt, Frankfurt, Germany*

ABSTRACT

The challenge in inferring cognitive processes from observational data is to correctly align overt behavior with its covert cognitive process. To improve our understanding of the overt–covert mapping in the domain of decision making, we collected eye-movement data during decisions between gamble-problems. Participants were either free to choose or instructed to use a specific choice strategy (maximizing expected value or a choice heuristic). We found large differences in looking patterns between free and instructed choices. Looking patterns provided no support for the common assumption that attention is equally distributed between outcomes and probabilities, even when participants were instructed to maximize expected value. Eye-movement data are to some extent ambiguous with respect to underlying cognitive processes. Copyright © 2017 John Wiley & Sons, Ltd.

KEY WORDS process tracing; expected value; priority heuristic; eye tracking; reverse inference

INTRODUCTION

The focus of research in the field of judgment and decision making has shifted from understanding *what* people choose to *how* they decide (Ford, Schmitt, Schechtman, Hults, & Doherty, 1989; Johnson & Ratcliff, 2014; Krajbich, Armel, & Rangel, 2010; Svenson, 1979, 1996). Various process tracing methods—including thinking aloud, mousetracking, eye tracking, and brain imaging—have been instrumental in this development (Schulte-Mecklenbeck, Kühberger, & Ranyard, 2011a, 2011b). In all process tracing methods, a close relationship between overt information acquisition and underlying covert cognitive processes is a core assumption. For example, activation of some brain area, measured with electroencephalography or functional magnetic resonance imaging, is taken to be indicative for cognitive processes associated with this area. The reasoning works as follows: (i) When a task recruits some psychological process *Pr*, brain activation pattern *A* is likely to be found; (ii) in the present study, a pattern of activation *A* was found when task *T* was presented; it can therefore be concluded that (iii) the psychological process *Pr* was recruited by task *T* (Machery, 2014). This reasoning is deductively invalid because different processes can be responsible for the same observable patterns. Poldrack (2006, p. 59) has diagnosed an "epidemic of reasoning" in the neuroscience literature, with researchers commonly committing the fallacy of *reverse inference*.

Can such reverse inferences also be found in decision-making studies? We think that this is the case and want to illustrate the problem by presenting two different perspectives on risky decision making—the perspective of an economist versus that of a psychologist. Imagine that both the economist and the psychologist are running an eye-tracking study, with participants making choices with gamble problems. Both observe a participant's gaze moving from an outcome (O) to the adjacent probability (P): $0 \rightarrow P$.[1] The economist might interpret this pattern within an *expected value* (EV; Hacking, 1984) framework where "plausible implications" for the processing of information are (i) information acquisition develops within and between gamble-problems; (ii) outcomes and probabilities are processed pairwise and consecutively; and (iii) each gamble receives the same amount of attention (Payne, Braunstein, & Carroll, 1978). The economist thus concludes that the $0 \rightarrow P$ pattern is indicative of calculating an EV. The psychologist, on the other hand, might prefer a heuristic framework. He favors, for instance, the priority heuristic (PH; Brandstätter, Gigerenzer, & Hertwig, 2006) as a model of risky choice. The $0 \rightarrow P$ pattern for him is indicating reading,[2] a pre-choice behavior necessary before the actual choice can be taken. That is, the two models force the two observers into radically different interpretations of the same behavioral pattern. Nothing in the $0 \rightarrow P$ transition, except its duration (given other assumptions about processing speed for calculation vs. comparison), can be used to identify the correct interpretation.

How prominent is this reasoning in the process tracing literature? Indeed, there are numerous examples where overt patterns, possibly indicative of some cognitive process, are

*Correspondence to: Michael Schulte-Mecklenbeck, University of Bern, Engehaldenstr. 4, 3012 Bern, Switzerland. E-mail research@schulte-mecklenbeck.com

[1]We denote → as a transition between two pieces of information, in this example, from an outcome (O) to a probability (P).
[2]Brandstätter, Gigerenzer, and Hertwig (2008, p. 286) dissect a first reading phase into reading all eight pieces of information available in a gamble—which would result in four outcome–probability transitions (and three other types of transitions not relevant here).

interpreted within some theoretical account, possibly falling prey to the reverse inference problem (Brandstätter et al., 2006; Day, 2010; Fiedler & Glöckner, 2012; Glöckner & Herbold, 2011; Johnson, Schulte-Mecklenbeck, & Willemsen, 2008; Pachur, Hertwig, Gigerenzer, & Brandstätter, 2013; Riedl, Brandstätter, & Roithmayr, 2008; Su et al., 2013).

In what follows, we propose an experimental approach that turns this direction of inference upside down; we emphasize the *forward inference* route from cognitive processes to behavioral patterns by making acquisition processes explicit through instruction. For instance, by instructing a participant to multiply an outcome with a probability, we expect at least one acquisition of each component (O and P) as well as at least one transition between the two (either $O \rightarrow P$ or $O \rightarrow P$). Failure to observe this pattern renders the mapping of overt behavior and cognitive processes questionable. The same reasoning applies for other mathematical operations (e.g., subtraction), or simple comparisons: If people were comparing two outcomes, we ought to see an $O \rightarrow O$ pattern. Comparing patterns following from instructed strategies to patterns where no instructions are provided thus enables an evaluation of the link between overt and covert processes (i.e., between looking and thinking).

The idea of instructing participants to follow specific strategies has been used in research on visual attention, where the effects of instruction (top-down) have been compared with those of stimulus-driven (bottom-up) processes. Yarbus (1967) was among the first to establish that different instructions (e.g., searching for a specific person vs. counting the number of people in a picture) result in different eye-movement patterns. His results have more recently been replicated for visual perception (DeAngelus & Pelz, 2009) and consumer choices (Glaholt, Wu, & Reingold, 2010; Pieters & Wedel, 2007).

For our interest in risky gamble problems, it is beneficial to use strategies that can be distinguished easily. We define a gamble problem as a choice between two gambles (A or B). Each gamble consisted of two outcome–probability pairs (see Appendix A for a full list of the gamble problems, which were taken from Pachur et al., 2013). We chose to pit two fundamentally different perspectives on risky choice against each other: (i) EV theory,[3] which consists of first calculating the EV for each gamble and then choosing the gamble with the maximum (i.e., higher) EV, and (ii) the PH (Brandstätter et al., 2006), which makes choices based on a comparison[4] of different gamble components: the minimum gain, the maximum gain, and their probabilities, respectively.

What are the key differences between these decision strategies? An EV is calculated for each gamble by summing up the products of all O–P pairs and then choosing the gamble offering the higher EV. The PH, in contrast, applies three rules (in this example, for gambles consisting of gains): (i) the *priority rule* is going through gamble problems such that the minimum gain (i.e., the minimum amount of money to win) is considered first, followed by the probability of the minimum gain, and finally the maximum gain (i.e., the maximum amount of money to win); (ii) the *stopping rule* requires evaluation to stop if the minimum gains differ by at least 1/10 (or more) of the maximum gain, otherwise to stop if probabilities differ by at least 1/10 (or more) on the probability scale; and the (iii) *decision rule*, choosing the gamble with the more attractive gain or probability.[5]

For the sake of comparison, we can also formulate EV in terms of the same set of rules. In EV, the *priority rule* gives no priority to any aspect of a gamble problem. Thus, all Os and Ps have equal priority, as have small or large values. The *stopping rule* requires multiplication of all O–P pairs and storing EVs until all O–P pairs have been processed; and the *decision rule* requires choosing the gamble with the highest EV and, if no maximum EV exists, choosing randomly.

In the present study, we instruct one group of participants to follow EV and another group to follow PH while choosing between gambles and record each group's eye movements. As a within-subject comparison, we also collect eye-movement data in a control condition without strategy-specific instructions. In comparing these three conditions, we evaluate the mapping between instructed thought processes and overt information acquisition behavior.

METHOD

Participants

Fifty students from the University of Salzburg participated in the study. Two participants were excluded from the data analysis because they could not be calibrated on the eye-tracking system. The remaining 48 participants (34 women) had a mean age of $M = 23.5$ years ($SD = 4.7$). All participants had normal or corrected-to-normal eyesight.

Apparatus

Eye-movement data were recorded using an EyeLink CL eye tracker (SR Research Ltd, Ontario, Canada) with a sampling rate of 1000 Hz. Eye movements were obtained from the right eye, with the participant's head stabilized on a chin and head rest. The distance between the eye and the monitor (Vision Master Pro 454, Iiyama, Tokyo, Japan) was held constant at 52 cm (20.5 in.). The screen had a diagonal of 21 in. and a resolution of 1024×768 pixels; the refresh rate was 120 Hz. Stimuli were presented in Courier New font (18 point); a single letter corresponded to about 0.4° of visual

[3]We chose EV because (i) it utilizes only multiplication and addition at its core and (ii) it can be understood as a representative for several integration or expectation models that lend themselves to a similar process interpretation, such as subjective expected utility theory (Friedman & Savage, 1948) and (cumulative) prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992).

[4]Note that, for example, for estimation of the 10% threshold in the PH calculations are needed.

[5]According to Brandstätter, Gigerenzer, and Hertwig (2008), the term "attractive" refers to the gamble with the higher gain and the lower probability of the minimum gain. Depending on how many priority steps have to be executed, a distinction is made between one-reason (decision based on minimum gain), two-reason (decision based on probability of minimum gain), and three-reason gambles (decision based on maximum gain). These labels provide both information about characteristics of the gambles and procedural information about how a gamble should be solved.

DOI: 10.1002/bdm

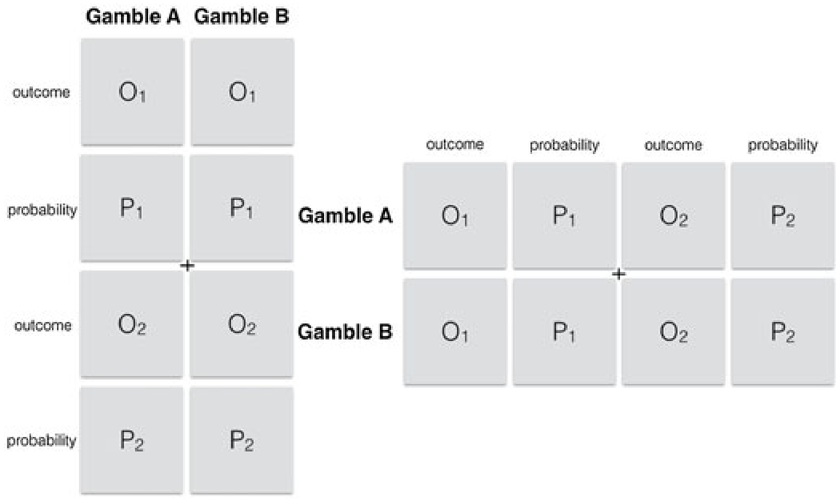

Figure 1. Two gamble problems, one in vertical (left) and one in horizontal (right) presentation format. Each gamble problem consists of two gambles (A and B), which in turn consist of two outcome–probability (O–P) pairs. The areas of interest (shown in gray and labeled O1, P1, …) and O–P labels are for presentation purposes only. A fixation cross was displayed before the onset of each trial

angle. The experiment was preceded by a nine-point calibration (using a grid of 3 horizontal positions × 3 vertical positions). After practice trials and before each experimental condition, a fixation check tested the accuracy of the measurements and initiated a full re-calibration in case of failure. Participants entered choices via a button box (Microsoft SideWinder Plug & Play Game Pad, Redmond, WA, USA).

Material

For each gamble problem, outcomes and probabilities were presented at equidistant locations in a 2 × 4 matrix, in either horizontal or vertical format (between subjects) to counteract reading effects (see Figure 1 for an example of each format). Participants were presented with 24 gamble problems in randomized order. Within each gamble problem, we counterbalanced the position of each outcome–probability pair in each left/right or top/bottom position.

Procedure

Participants first received general information about the eye-tracking setup, a description of the gamble problems, and four unrelated gamble problems as a warm-up task. All participants were then presented with the 24 gamble problems and were asked to choose their preferred gamble, without specific instructions. This will be referred to as the *no instruction* (NO) condition. Subsequently, the same 24 gamble problems were presented again, with half of the participants being instructed to apply the EV strategy and the other half to apply the PH. Participants were instructed in detail to do each step necessary to proceed according to the respective strategy, EV or PH. In addition, the relevant instructions were provided in paper form throughout the experiment. Instructions were such that, for every necessary step, the instructions identified (i) the target variable(s), for example, *outcome 1 of gamble A*; and (ii) the required process, for example, *look at*, *multiply*, *add*, *store in memory*, and *compare*. Appendix B provides translations of the German instructions, separately for EV and PH.

Participants received a flat payment of €5 and could earn a bonus of up to €5 based on their choices. We additionally raffled a €100 Amazon gift certificate among all participants. The whole experiment took approximately 45 minutes.

Data preparation

Eye-movement data were analyzed with a customized Perl script available from the third author. Saccade and fixation events were defined online by the host of the eye-tracking system and read out by the script if a fixation landed on one of the pre-defined, non-overlapping areas of interest (AOIs). Eight AOIs with a size of 60 × 60 pixels were defined, with the numeric values in the center. Fixation durations shorter than 50 milliseconds were removed from the analysis. The number and duration of fixations were analyzed for each AOI. All analyses were run with R version 3.2.4 (R., 2016).[6]

Predictions

Our analysis will focus on the following process measures: (i) acquisition patterns, (ii) frequencies of acquisitions, and (iii) durations of acquisitions. For each of these measures, we will derive a separate set of predictions. Table 1 provides a detailed set of predictions (and results) for fixation frequencies (also specifying expected differences for one-, two- and three-reason gambles).

Acquisition patterns

In EV, we instructed participants to base their decision on the calculation and comparison of an EV, for example, to

[6]Data and code are available at https://github.com/michaelschulte/InducingCognitiveProcesses.

DOI: 10.1002/bdm

Table 1. Comparisons between outcomes and probabilities with predictions of acquisition patterns separately for EV and PH

	Predictions		Empirical findings			RMSD		
Comparison	EV	PH	EV	PH	NO	NO–EV		NO–PH
O vs. P	$O_{r=1} = P_{r=1}$	$O_{r=1} > P_{r=1}$	57.6 > 42.4	77.7 > 22.3	56.2 > 43.8	12.8	<	27.1
	$O_{r=3} = P_{r=3}$	$O_{r=3} > P_{r=3}$	54.3 > 45.7	58.8 > 41.2	48.9 < 51.1	12.1	<	17.4
O^{max} vs. O^{min}	$O^{max}_{r=1} = O^{min}_{r=1}$	$O^{max}_{r=1} < O^{min}_{r=1}$	64.7 > 56.0	82.8 > 76.0	62.8 > 55.6	18.8	<	30.2
	$O^{max}_{r=2} = O^{min}_{r=2}$	$O^{max}_{r=2} < O^{min}_{r=2}$	68.6 > 43.9	83.6 > 47.8	62.9 > 44.3	16.4	<	22.9
P^{max} vs. P^{min}	$P^{max}_{r=2} = P^{min}_{r=2}$	$P^{max}_{r=2} < P^{min}_{r=2}$	31.4 < 56.1	16.4 < 52.2	37.1 < 55.7	16.4	<	22.9
	$P^{max}_{r=3} = P^{min}_{r=3}$	$P^{max}_{r=3} < P^{min}_{r=3}$	44.0 < 53.0	23.7 < 56.0	48.2 < 56.9	20.0	<	27.9
$O^{max}_{r=1}$, $O^{max}_{r=2}$ vs. $O^{max}_{r=3}$	$O^{max}_{r=1} = O^{max}_{r=3}$	$O^{max}_{r=1} < O^{max}_{r=3}$	64.7 > 56.0	82.8 > 76.3	62.8 > 51.7	18.5	<	35.0
	$O^{max}_{r=2} = O^{max}_{r=3}$	$O^{max}_{r=2} < O^{max}_{r=3}$	68.6 > 56.0	83.6 > 76.3	62.9 > 51.7	16.5	<	35.0
$P^{min}_{r=1}$ vs. $P^{min}_{r=2}$, $P^{min}_{r=3}$	$P^{min}_{r=1} = P^{min}_{r=2}$	$P^{min}_{r=1} < P^{min}_{r=2}$	44.1 < 56.1	42.9 < 52.2	44.4 < 55.7	19.1	<	28.1
	$P^{min}_{r=1} = P^{min}_{r=3}$	$P^{min}_{r=1} < P^{min}_{r=3}$	44.1 < 53.0	23.9 < 56.0	44.4 < 56.9	19.1	<	28.1

Note: Empirical findings are shown as percentages of acquisitions per gamble problem (e.g., O vs. P). RMSDs are shown between the NO condition and the EV/PH conditions, respectively. Differences between acquisition frequencies are significant, with all ts > 3.7, ps $< .001$, but note the interdependence of these tests. RMSD, root mean square deviations; O, outcomes; P, probabilities; max, min, maximum or minimum values; $r = 1$, $r = 2$, $r = 3$: one-, two-, or three-reason gambles.

multiply O and P. Hence, we expect frequent within-gamble transitions $O \rightarrow P$ or $P \rightarrow O$. In PH, we instructed participants to base their decisions on the PH, for example, to compare minimum outcomes. Hence, we expect frequent $O \rightarrow O$ or $P \rightarrow P$ transitions between gambles, depending on the number of comparisons.

Frequency of fixations

In EV, we expect no difference in the number of fixations of O and P.

In PH, the number of fixations to O and P is dependent on the number of reasons (one-, two-, or three-reason gambles; Brandstätter & Körner, 2014).

Speculatively, we would also expect that execution of PH results in more fixations than EV given its rather complex sequential nature.[7]

Duration of fixations

The fixation duration—that is, the duration between two saccades in which the eye rests on a specific location—is taken to be indicative of attention (Rayner, 1998). However, it may also indicate consumption of cognitive resources. For instance, if it is more effortful to multiply an outcome by a probability (EV) than to compare the size of two outcomes (PH), the average fixation duration should be longer for multiplication than for comparison. It is clear from the literature that duration of fixation is associated with the complexity of the process executed (Velichkovsky, 1999). Hence, EV, which involves both multiplication and adding processes, can be expected to result in longer fixation times than PH, which requires only simple comparisons.

RESULTS

Choices

First, we evaluated whether participants' choices were consistent with the strategy they were instructed to use: EV or PH. We restricted this choice analysis to gamble problems 3, 6, 9, 11, 14, 17, 20, and 23 (Appendix A), because in these problems, (i) the EV ratio was not 1, making it possible to decide based on the EV, and (ii) EV and PH make different predictions. The instructions were effective in guiding participants to the predicted choices. Participants instructed to use EV chose the higher EV gamble in 62% of the gambles; participants instructed to use the PH made corresponding choices in 80% of the gambles. In the NO condition, 58% of choices were in line with EV, and 42% in line with PH. Note that it is difficult to choose correctly in the EV condition, because the difference in EV is of a maximum of only 7% (see the ratios of EVs for the gamble problems listed in Table A1). In a pilot study for this experiment, we used a gamble set with less difficult gambles in a similar experimental paradigm. Here, participants in the EV condition made 72% corresponding choices, while participants in the PH condition made 79% corresponding choices.[8]

Acquisition patterns

The EV and PH differed substantially in terms of the sequence in which information was accessed. We analyzed these patterns by calculating an overall search metric (SM).

The search metric

A popular method for characterizing overall patterns of information acquisition is to calculate the ratio of within- to between-gamble transitions by applying Böckenholt's (Böckenholt & Hynan, 1994) *SM* calculation:

[7] Counting the number of words in the two instructions (EV and PH) illustrates this expectation: We needed 581 words to instruct participants to use EV and 710 words to use PH. A translation of the original instructions (in German) to English is available in Appendix B.

[8] Note that for all remaining analyses, the full gamble problem set (Appendix A) was used.

DOI: 10.1002/bdm

$$SM = \frac{\sqrt{N}\left(\left(\frac{A*G}{N}\right)(WG - BG) - (G - A)\right)}{\sqrt{A^2(G-1) + G^2(A-1)}} \quad (1)$$

This SM distinguishes *within-gamble* transitions (e.g., the outcome and the probability of a single gamble are fixated in succession) from *between-gamble* transitions (e.g., two outcomes, or two probabilities, of different gambles are fixated in succession). The stimulus setup is represented in terms of the number of gambles G (two in our experiment) and the number of attributes A (four in our experiment); N denotes the total number of transitions. We replaced the absolute occurrences of transitions N with proportions to capture the index's sensitivity to large Ns (see Pachur et al., 2013; Schulte-Mecklenbeck, Sohn, de Bellis, Martin, & Hertwig, 2013, for a similar approach). An $SM > 0$ indicates a predominance of within-gamble transitions (expected for EV); an $SM < 0$ indicates a predominance of between-gamble transitions (expected for PH).

From the three instructed strategies, EV had, on average, the highest, positive index (standard deviations in parenthesis), $SM_{\mathrm{EV}} = 7.4$ (3.0) indicating strong within-gamble search; PH resulted in a negative index of $SM_{\mathrm{PH}} = -2.1$ (5.0), indicating between-gamble search; finally, the NO condition resulted in a positive index, $SM_{\mathrm{NO}} = 3.2$ (4.3; Figure 2). Note that the NO condition on average produced a positive SM when we compared EV and PH, significantly different from the other two, instructed conditions.

These descriptive findings were supported statistically by a multilevel regression with "participants" and "gamble problem" as random intercepts and "condition" (EV, PH, and NO) as a fixed effect. The SM index indeed was lower in NO than in EV, $b = -4.0$, $\mathrm{CI}_{95\%} = [-4.4, -3.6]$ and higher in NO than in PH, $b = 5.1$, $\mathrm{CI}_{95\%} = [4.7, 5.5]$.

Fixation frequencies

For the analysis of fixation frequencies, we first evaluated the three conditions on an aggregate level and then turned to a more fine-grained analysis of acquisition behavior.

The average number of fixations was highest in EV with $Fix_{\mathrm{EV}} = 91.6$ (24.3), indicating that a comparably large number of fixations was needed to follow this strategy. Indeed, the eight AOIs were fixated up to 200 times by several participants. The second instructed strategy, PH, resulted in $Fix_{\mathrm{PH}} = 69.3$ (9.2). Not surprisingly, participants used the fewest number of fixations when not provided with instruction (NO), $Fix_{\mathrm{NO}} = 34.9$ (6.4; Figure 3).

We followed up with a multilevel regression analysis with participants and gamble problem as random intercepts and condition (EV, PH, and NO) as a fixed effect. This analysis confirmed that the average number of fixations was lowest in NO than in both EV, $b = -58.4$, $\mathrm{CI}_{95\%} = [-61.9, -55.0]$, and PH, $b = -32.7$, $\mathrm{CI}_{95\%} = [-36.2, -29.3]$. No difference between EV and PH was found.

Next, we conducted a fine-grained analysis of acquisitions of outcomes and probabilities, as proposed by Johnson et al. (2008), who derived a list of predictions from the steps that Brandstätter et al. (2006) identified as necessary in applying the PH (Pachur et al., 2013, tested the same set of predictions). This analysis takes the relative size of outcomes and probabilities into account by distinguishing between the minimum outcome, the probability of the minimum outcome, the maximum outcome, and the probability of the maximum outcome. Table 1 presents the 10 predictions derived for EV and PH and the corresponding empirical findings from our experiment. The first line, for example, shows predictions for outcomes and probabilities (O vs. P). For one-reason gambles ($r = 1$), EV predicts equal numbers of acquisitions of outcomes and probabilities, as indicated by $O_{r=1} = P_{r=1}$. In contrast, PH predicts more acquisitions of outcomes than of probabilities, as indicated by $O_{r=1} > P_{r=1}$. We report the percentages of acquisition of outcomes and probabilities separately for the three conditions.

Two points are noteworthy: (i) EV always predicts an equal distribution of acquisition frequencies, regardless of the type of information (O or P) attended to, or the number of reasons (one, two, or three) considered (see priority rule for EV above). (ii) PH predicts more acquisitions of outcomes than of probabilities, and fewer acquisitions of the maximum outcome than of the minimum one. Inspection of

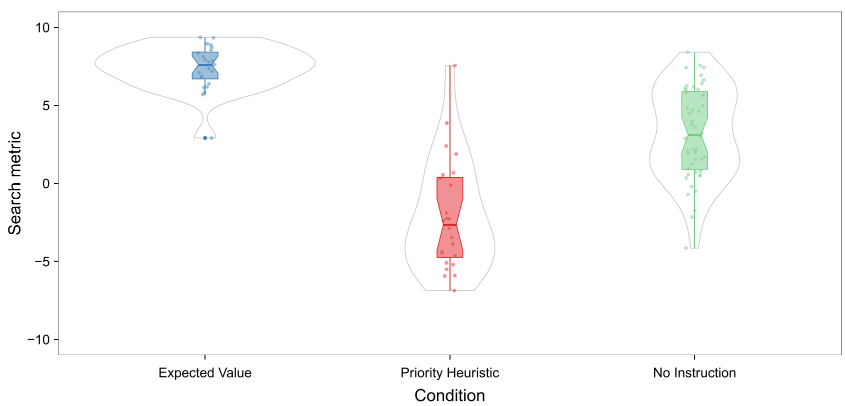

Figure 2. Search metric depicted as a box plot for each of the three conditions, with superimposed averaged raw data for each participant (jittered) and probability density function. Positive values indicate within-gamble transitions; negative values indicate between-gamble transitions

DOI: 10.1002/bdm

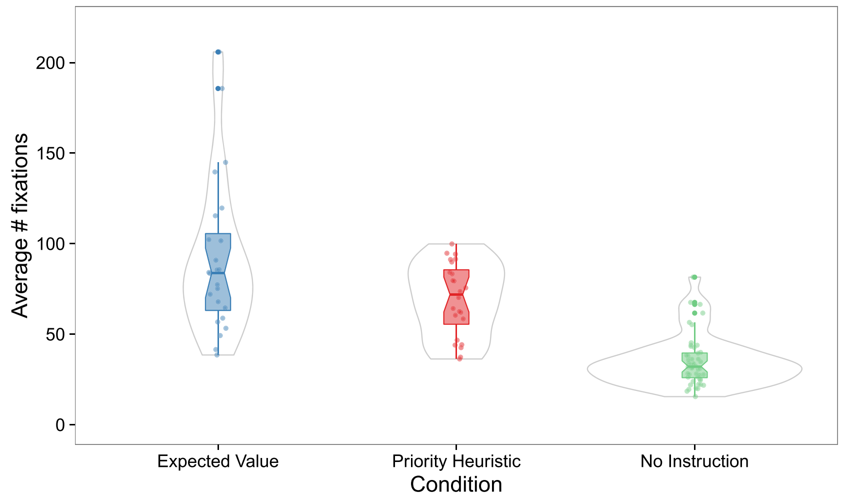

Figure 3. Fixation frequencies depicted as a box plot for each of the three conditions, with superimposed averaged raw data for each participant (jittered) and probability density function

Table 1 shows that, overall, outcomes were fixated more frequently than probabilities, replicating the finding of Pachur et al. (2013). In addition, larger outcomes attracted more fixations than smaller ones. The pattern for probabilities was the opposite: Smaller probabilities attracted more fixations than larger ones.

Assuming that the number of fixations is a proxy for importance (Velichkovsky, 1999) or weight (Schkade & Johnson, 1989; Wedell & Senter, 1997; Willemsen, Böckenholt, & Johnson, 2011), the fixation frequencies tell a simple story: Outcomes are more important than probabilities, and larger outcomes are more important than smaller ones. Instructing participants to use specific strategies seems to have little effect on this basic pattern. This indicates a bottom-up effect (e.g., of features of the gamble problems) that is stronger than the top-down effect of instructions. These findings are also inconsistent with the assumption that attention is distributed equally between outcomes and probabilities when calculating an EV, as all differences in the EV condition were significant.

Our method of inducing thought processes allows another comparison that can shed light on information acquisition. Specifically, we determined the difference between the NO condition and the two instruction conditions by calculating the root mean square deviation (RMSD; Hyndman & Koehler, 2006) of acquisitions between the NO condition and the EV/PH condition, respectively.

$$RMSD = \sqrt{\frac{\sum_{t=1}^{n}\left(x_{NO,\ t} - x_{[EV,\ PH],t}\right)^2}{n}} \quad (2)$$

To this end, the RMSD was calculated for each comparison presented in Table 1 (e.g., O vs. P between the NO and EV/PH conditions, respectively) across all $n = 24$ gamble problems and then averaged across participants. A smaller RMSD indicates greater similarity between two comparisons.

For all comparisons listed in Table 1, the RMSDs were smaller for the NO–EV comparison than for the NO–PH comparison. This means that the association between NO and EV, in terms of acquisition frequencies, is closer than that between NO and PH. We now turn to another measure derived from eye-movement behavior: fixation duration.

Fixation duration

The average length of fixations was highest in EV with $Length_{\mathrm{EV}} = 254.1$ milliseconds (11.4) followed by PH, with $Length_{\mathrm{PH}} = 220.0$ milliseconds (10.8). The shortest fixation durations were found for NO, $Length_{\mathrm{NO}} = 207.7$ milliseconds (10.3; Figure 4).

As expected, we found a main effect of condition in a multilevel regression with participants and gamble problem as random intercepts and condition (EV, PH, and NO) as a fixed effect. The fixation duration for the NO condition was shorter than EV, $b = -40.5$, $\mathrm{CI}_{95\%} = [-43.7, -37.4]$, and PH, $b = -18.1$, $\mathrm{CI}_{95\%} = [-21.3, -15.0]$.

Let us briefly sum up our results, before turning to a more detailed discussion of the issues at hand. The clearest results are provided with the acquisition patterns (refer to section on Acquisition Patterns) and the detailed fixation frequency analysis (Table 1). The positive values of the SM for EV and NO indicate, on an aggregate level, more transitions within gambles than between gambles. The RMSD, describing the distance between two measures, is smaller for the comparison NO–EV than for NO–PH in all of the 10 suggested tests, indicating a closer relationship between NO–EV than NO–PH. For the overall fixation frequency and fixation duration, EV and PH are more closely related and different to NO; for these measures, it is less clear which strategy is followed. These patterns indicate that both instructed strategies (EV and PH) capture some, but not all, of the processes used by the participants in the NO condition.

DISCUSSION

In this study, we instructed two decision strategies, EV and PH, and investigated four dependent measures (choices,

DOI: 10.1002/bdm

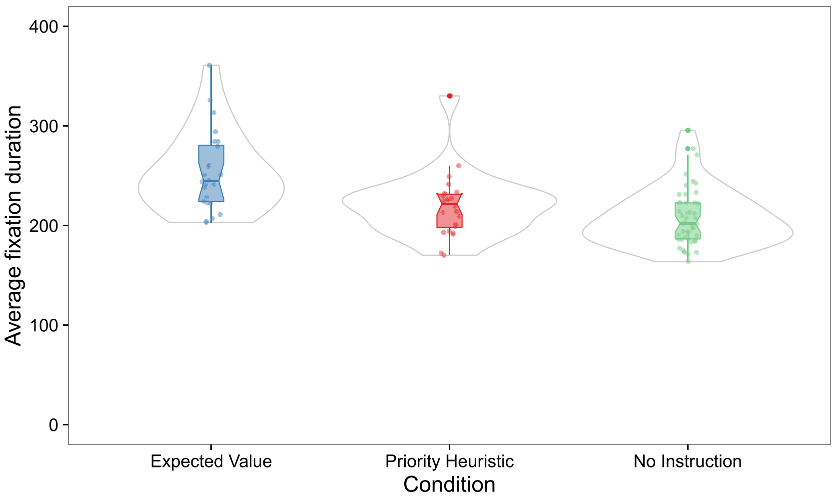

Figure 4. Average fixation depicted as a box plot for each of the three conditions, with superimposed averaged raw data for each participant (jittered) and probability density function

acquisition patterns, acquisition frequencies, and fixation duration) to better understand how strategy use is reflected in these measures. What do we learn from our results? For a differentiation between the two instructed strategies (EV and PH) and the NO condition, the acquisition patterns as well as the fixation data speak a clear language: Without instruction, participants' looking patterns indicate strategies more closely related to EV than to PH. This interpretation rests on four supporting facts: (i) Our findings are built on detailed, model derived, predictions, and (ii) are consistent with work by Johnson et al. (2008), as well as with work on the distribution of attention on probabilities and outcomes (Glöckner & Herbold, 2011; Huber, Wider, & Huber, 1997; Pachur et al., 2013; Su et al., 2013). (iii) Our results hold across different dependent variables, that is, SM and frequencies. (iv) The within-subject comparison of patterns generated in NO compared with instructed EV or PH strategies provides a unique perspective on the mapping between cognitive processes and overt behavior.

We will now turn to a more detailed evaluation of our data.

Choice and fixation patterns depend on gamble difficulty

Why was the correspondence between choices in the EV condition and the number of EV-maximizing choices so low? At least four reasons can explain this result: Participants (i) were unable, (ii) were unwilling to follow the instructions, (iii) used a different strategy, or (iv) committed calculation errors. We are confident that reasons (i) and (ii) did not generally apply in our study out of the following reasons: We took great care in formulating and testing our instructions (in a pilot study); we provided extensive written information on how to calculate EVs in terms of algebraic steps; we repeated the written information verbally; we made participants go through an example gamble to make sure they made the correct calculation—this step was repeated until all participants were able to calculate EVs correctly; we provided a diagram version of the instructions in both the EV and PH conditions for participants throughout the experiment. To ensure high motivation to complete the task correctly, we also offered financial incentives for correct choices, as is common practice in the field of economics.

Of course, we cannot rule out that simplified versions of EV and PH were used when participants were instructed to follow these rules. Inspecting the SM, the number of fixations, and the length of fixations, it seems unlikely that this was the case as the results match the theoretically assumed patterns closely (high SM index, more fixations than PH/NO, and more fixation duration than PH/NO).

For instance, simplifying calculations, especially in difficult choices, might be tempting. The difficulty of gambles is often gauged in terms of the ratios of the gambles' EVs: Difficult gambles have ratios equal or close to 1; easy gambles have ratios sufficiently different from 1. The average ratio for the gamble set we used was 1.01 ($SD = .03$), indicating that our gambles were rather difficult. Pachur et al. (2013) showed that the fit of EV to actual choices is dependent on gamble difficulty, with 75% correct predictions for easy gambles but only 38% for difficult gambles. This result might explain our comparably low matches to the EV predictions.

Against this background, we inspected the gamble set used in the present study (from Pachur et al., 2013) in more detail. It appears that the number of reasons (Appendix A) and the difficulty of EV calculations are confounded in this set. To illustrate this point, we calculated the average number of outcome digits[9] for each gamble. The one-reason gambles had, on average, 16.7 digits; the two-reason gambles, 15.0 digits; and the three-reason gambles, 12.6 digits. Hence, as the number of reasons increases, gambles become numerically easier to calculate and easier to compare,[10] a pattern that is also reflected in the decrease in fixation duration when such calculations have to be performed. In the mathematical cognition literature, such issues are discussed under the heading "problem-size effect." This effect describes the decrease

[9]For example, an outcome of 1000 has four digits. We focus on outcomes because there is less variation in the number of digits in probabilities, which are constrained in the interval [0, 1].

[10]Correlating the number of digits with the number of reasons, we found a negative relationship of $r(2302) = -.55$, $p < .001$.

DOI: 10.1002/bdm

in completion time of a calculation as the number of digits decreases (Zbrodoff & Logan, 2004). It is interesting to note that most problem sets used in this literature are much easier (i.e., have fewer digits) than those used in studies utilizing gamble problems (DeWolf, Grounds, Bassok, & Holyoak, 2014; Raghubar, Barnes, & Hecht, 2010).

Another feature of the gamble-problem set is the number of zero outcomes. Of the 6 two-reason gamble problems, 5 gambles have zero outcomes (20.8% of the 24 values); of the 6 three-reason gamble problems, 8 gambles have zero outcomes (25%; Table 2). In terms of calculating an EV, a zero value is a game changer. Whereas a standard gamble problem requires four multiplications, each zero outcome removes one of those, making it increasingly easy to calculate or compare. Note also the ambiguity of the value of zero for process measures: Would we expect an O–P pattern if the outcome is zero? Probably not, as in this case no multiplication is needed; the EV strategy thus simplifies to 0. Zero outcomes thus might be responsible for our finding that outcomes are inspected more often than probabilities, as well as for the decrease in fixation duration across reasons in the EV condition. This shows the yet overlooked importance of such bottom-up features of the problem sets.

Interestingly, our data did not support the prediction of an increase in processing time between one-, two-, and three-reason gambles in the PH. As with the increased number of reasons, more comparisons are necessary for each gamble problem; hence, longer processing/reaction times should result (Brandstätter et al., 2006). In our positive outcome gamble set, a zero value is always the minimum outcome. According to PH's priority rule, the minimum is the most important value for a choice. It remains a challenging task to decide whether the special role of the zero outcome is derived from its calculation consequences or from its processing priority. Overall, based on these findings, we conclude that choices may not closely conform to the instructions unless (i) the outcomes are sufficiently small to enable correct unaided mathematical operations; (ii) only simple operations are required; or (iii) zero outcomes change the processing strategy. In decision research, these conditions are frequently violated.

Model comparisons and the merits of inducing thought processes

Our set of directional tests, utilizing fixation frequencies on outcomes and probabilities, demonstrates the value of inducing thought processes. In the literature, much use is made of a "scorekeeping heuristic"—in other words, the results of a process tracing study are classified pro or contra a prediction and tallied (cf. Johnson et al., 2008). Applying this heuristic, we can make two interesting comparisons: comparing the two instructed strategies to theoretical predictions and comparing them to the actual choices in the NO condition. Concerning the theoretical comparison for EV, Table 1 shows that for 10 out of 10 tests, the theory predicts wrongly for outcomes and probabilities (see Table 1, columns 1 vs. 3). For PH a score with 6 out of 10, directional predictions being supported point at a performance close to chance level. The second perspective, comparing the two instructed strategies to the NO shows a similar pattern—the equality predictions of EV cannot be found in the associated NO condition (in Table 1 compare columns 1 and 5). Comparing the NO condition to the PH predictions would also be inconclusive as five tests speak for, and five against, the patterns predicted by PH (see columns 2 and 5).

Adding the instructed EV and PH conditions changes this evaluation. This pattern leads to the rejection of the hypothesis that participants actually used EV-related processes. Furthermore, our data are clearly not consistent with the assumption of an equal distribution of attention in EV. Quantifying the differences between NO and the two instructed strategies reveals a closer relationship between NO and EV than between NO and PH—a conclusion that could not otherwise have been drawn. Finally, the striking similarity in the directionality of results for the two instructed and the NO condition points to an interaction with factors other than the instructions alone. These findings thus indicate an issue largely ignored in the literature on judgment and decision making: The bottom-up effects of how stimuli are constructed (Pleskac & Hertwig, 2014).

Our method of inducing thought processes was informed by the lively discussion on how cognition is, and should be, studied in the field of neuroscience (Henson, 2006; Hutzler, 2014; Poldrack, 2006). Specifically, we were interested in the common, but deductively invalid, practice of drawing inferences, from observed brain activations, about the cognitive processes recruited by specific tasks (reverse inference). What are the commonalities and differences between reverse inference and inducing thought processes? Of course, in our example of choice behavior, there is no mapping to a specific brain area, as would be the case in a neuroscience study. However, the observation and interpretation of a fixation, or fixation pattern, as indication of some cognitive process (e.g., attention or processing) seems to be closely related to the discussion in neuroscience.

Our approach builds on the concept of forward inference (Henson, 2006), which allows the predictions of different theories to be pitted against each other. To apply forward inference, researchers need to design experimental conditions that differ in the light of one theory, but not the other, following a dissociation logic (Teuber, 1955). In neuroscience, this procedure allows assigning an observed activation to the respective theory and the corresponding cognitive process. We believe that inducing thought processes by instruction can generate similar dissociations—in our example, between two competing choice models, but also more generally, between competing concepts about cognition. Our findings lend themselves to forward inference to some degree, but it is clear that the to-be-compared models need much more fine-grained description when it comes to using data with the resolution generated by contemporary eye-tracking technology. Returning to our multiplication example: It is not enough to predict an $O \rightarrow P$ transition for multiplication, as the exact pattern depends on subtle features of the required computations, like the size of the numbers, the number of digits, or the number of zeros. Presumably, the eyes will fixate differently when doing the 6×0.3 multiplication and storing its result for later comparison, than when doing the same with

DOI: 10.1002/bdm

1650 × 0.85. In addition, is the difference in magnitude best accounted for by the frequency of fixations (larger magnitude—more fixations), or by the duration of fixations (larger magnitude—longer fixation)? Some of these questions can be answered in research using eye-tracking technology, but concurrent collection of other types of process data, being conceptually closer to the studied process, for example, thinking aloud protocols (Russo, 2011), might be necessary to gain a thorough understanding of the relevant processes.

Tracking eye-movement results in minute details, and forward inference accordingly requires precise models that incorporate these details. Other disciplines have put forward models that capitalize on the high-resolution data available. For example, one of the prominent reading models, the EZ Reader (Reichle, Rayner, & Pollatsek, 1999), makes detailed predictions on reading behavior for the bottom-up level (e.g., oculomotor control and visual processing) and the top-down level (e.g., attention).

Conclusions

Our experiment demonstrates that inducing thought processes is a useful tool to (i) identify incorrect assumptions about the distribution of attention (e.g., that attention is equally distributed between outcomes and probabilities when people calculate EVs) and (ii) evaluate the process predictions of competing models. A key factor in achieving the second point is that researchers collect data on how participants actually acquire information when they apply a specific strategy rather than predicting how they should acquire information based on purely theoretical considerations. Theories should move toward precise predictions acknowledging the level of detail provided by modern process tracing technology. Using current process tracing technology uncovers the necessity of developing more detailed models of decision making. In Marr's (1982) terminology, describing models at the computational level, where problems are specified in the generic manner that is typical for decision theory, is not enough. We need to formulate our models also on the algorithmic level, describing how exactly the computational problems are solved. This is the bridge to the implementational level, which identifies the neuronal mechanisms and their organization actually performing the computation. One of the priority heuristic's strengths is the detailed process predictions on an algorithmic level. With eye-movement technology, correct forward inference (i.e., predicting eye-movement patterns from strategy) and correct reverse inference (i.e., inferring strategy use from eye-movement patterns) require a close fit between the graininess of data and the graininess of theory. Somewhat paradoxically, the resolution of data from eye-tracking seem to be "too high" for the models tested here, namely, EV and PH. Put differently, decision models need to be fleshed out in much more processing details instead of paramorphic as-if-descriptions in order to lend themselves to the decisive test by eye-movement technology.

ACKNOWLEDGEMENTS

We thank the mighty Susannah Goss for editing the manuscript. Sabine Bitter and Isabella Mühlgrabner helped with the data collection. We are grateful to Eduard Brandstätter, Andreas Glöckner, Ralph Hertwig, Thorsten Pachur, and Timothy J. Pleskac for comments on an earlier version of this manuscript.

APPENDIX A:

Table A1. The 24 gamble problems presented

	Gamble A				Gamble B						
	Option 1		Option 2		Option 1		Option 2				
#	O1	P1	O2	P2	O1	P1	O2	P2	EV ratio	Source	Reasons
1	2000	0.6	500	0.4	2000	0.4	1000	0.6	1	BGH	1
2	4000	0.2	2000	0.8	3000	0.7	1000	0.3	1	BGH	1
3	800	0.8	500	0.2	820	0.6	600	0.4	1.01	PHGB	1
4	5000	0.7	100	0.3	5000	0.65	1000	0.35	0.98	PHGB	1
5	−500	0.4	−2000	0.6	−1000	0.6	−2000	0.4	1	BGH	1
6	−2000	0.7	−5000	0.3	−2800	0.9	−4800	0.1	0.96	PHGB	1
7	−500	0.2	−800	0.8	−600	0.4	−820	0.6	1.01	PHGB	1
8	−50	0.3	−3500	0.7	−600	0.35	−3400	0.65	1.02	PHGB	1
9	−100	0.3	−5000	0.7	−1000	0.35	−5000	0.65	0.98	PHGB	1
10	−50	0.1	−900	0.9	−400	0.15	−880	0.85	1.01	PHGB	1
11	−2000	0.6	−2350	0.4	−1700	0.55	−2500	0.45	1.04	PHGB	1
12	−150	0.7	−2500	0.3	−650	0.9	−2400	0.1	1.04	PHGB	1
13	2000	0.5	0	0.5	4000	0.2	300	0.8	0.96	PHGB	2
14	1600	0.3	1000	0.7	1300	0.5	1000	0.5	1.03	PHGB	2

(*Continues*)

DOI: 10.1002/bdm

Table A1. (Continued)

	Gamble A				Gamble B						
	Option 1		Option 2		Option 1		Option 2				
#	O1	P1	O2	P2	O1	P1	O2	P2	EV ratio	Source	Reasons
15	1800	0.2	200	0.8	1000	0.4	200	0.6	1	PHGB	2
16	6000	0.45	0	0.55	3000	0.9	0	0.1	1	KT	2
17	−1000	0.7	−1600	0.3	−1000	0.5	−1300	0.5	1.03	PHGB	2
18	0	0.55	−6000	0.45	0	0.1	−3000	0.9	1	KT	2
19	6000	0.3	2500	0.7	8200	0.25	2000	0.75	1	BGH	3
20	3000	0.4	2000	0.6	3600	0.35	1750	0.65	1.00	BGH	3
21	6000	0.001	0	0.999	3000	0.002	0	0.998	1	KT	3
22	4000	0.2	0	0.8	3000	0.25	0	0.75	1.07	KT	3
23	0	0.8	−4000	0.2	0	0.75	−3000	0.25	1.07	KT	3
24	0	0.999	−6000	0.001	0	0.998	−3000	0.002	1	KT	3

Note: Subscripts indicate corresponding O–P pairs. The three sources of our gamble set are BGH = Brandstätter et al. (2006), PHGB = Pachur et al. (2013), and KT = Kahneman and Tversky (1979). Following Brandstätter et al. (2006), 1–12 are one-reason gambles, 13–18 are two-reason gambles, and 19–24 are three-reason gambles.
O, outcomes; P, probabilities.

APPENDIX B:

Dear Participant,

Thanks a lot for participating in our experiment. You will now see information about different gamble problems on the computer screen.

These gambles are characterized through

(a) consequences (i.e., amounts of money to win or lose) and
(b) probabilities for each consequence.

Example:

Gamble A:	You win	20 with a probability of .1 and 1 with a probability of .9
Gamble B:	You win	4 with a probability of .4 and 3 with a probability of .6

Consequences can come as wins or losses, e.g., 20 or −20. Each of the consequences varies within a probability of 0 to 1. A probability of .05 means that a consequence will happen with a 5% probability; a probability of .95 means that a consequence happens with a 95% probability. Within one gamble (e.g., Gamble A), the sum of the probabilities of the two consequences is always 1.

Your task it to determine which of the two gambles has the higher expected value. The expected value of gambles can be determined through weighting the consequences by their probabilities.

Here is an example:

For *Gamble A* in the example above, the expected value is calculated as follows:

$$20\times0.1 + 1\times0.9 = 2 + 0.9 = \mathbf{2.9}$$

For *Gamble B* in the example above, the expected value is calculated as follows:

$$4\times0.4 + 3\times0.6 = 1.6 + 1.8 = \mathbf{3.4}$$

Hence, Gamble B has a higher expected value (3.4) than Gamble A (2.9).

For gamble problems where money can be lost, the following applies:

Example for losses:

Gamble A:	You lose	20 with a probability of .1 and 1 with a probability of .9
Gamble B:	You lose	4 with a probability of .4 and 3 with a probability of .6

For *Gamble A* in the example above, the expected value is calculated as follows:

$$-20\times0.1 + (-1)\times0.9 = -2 + (-0.9) = \mathbf{-2.9}$$

For *Gamble B* in the example above, the expected value is calculated as follows:

$$-4\times0.4 + (-3)\times0.6 = -1.6 + (-1.8) = \mathbf{-3.4}$$

Hence, Gamble A has a higher expected value (−2.9) than Gamble B (−3.4) because the expected loss is smaller.

For every correct answer, you receive 20 cents. You can make another EUR 4.80 by answering all gambles correctly. The amount you made will be paid to you in addition to your showup fee.

Please note that aids like pocket calculators or paper are not allowed during the study.

Exercise

We will now walk you through a step-by-step example of calculating the expected value:

DOI: 10.1002/bdm

Step 1:	Read each of the values once.
Step 2:	In Gamble A, multiply Value 1 with Chance 1 and remember the result. **10 × 50% = 5**
Step 3:	In Gamble A, multiply Value 2 with Chance 2 and remember the result. **0 × 50% = 0**
Step 4:	Add up the results from step 2 and 3 and remember the result. **5 + 0 = 5**
Step 5:	In Gamble B, multiply Value 1 with Chance 1 and remember the result. **20 × 20% = 4**
Step 6:	In Gamble B, multiply Value 2 with Chance 2 and remember the result. **1 × 80% = .8**
Step 7:	Add up the results from step 5 and 6 and remember the result. **4 + .8 = 4.8**
Step 8:	Choose the Gamble with the higher sum. **5 (Gamble A) > 4.8 (Gamble B) ➔ Gamble A ➔ Press A**

Dear Participant,

Thanks a lot for participating in our experiment. You will now see information about different gamble problems on the computer screen.

These gambles are characterized through

(a) consequences (i.e., amounts of money to win or lose) and
(b) probabilities for each consequence.

Example for gains

Gamble A:	You win	20 with a probability of .1 and 1 with a probability of .9
Gamble B:	You win	4 with a probability of .4 and 3 with a probability of .6

Consequences can come as wins or losses, e.g., 20 or −20. Each of the consequences varies within a probability of 0 to 1. A probability of .05 means that a consequence will happen with a 5% probability; a probability of .95 means that a consequence happens with a 95% probability. Within one gamble (e.g., Gamble A), the sum of the probabilities of the two consequences is always 1.

Your task is to determine which of the two gambles (A or B) should be chosen based on a **decision rule**. The decision rule consists out of one, two, or three steps. How many steps are needed to make a decision depends on the gamble problem at hand.

The rules used are different for gains and losses—we will explain each of these now in turns:

For **gains**, the following steps have to be used:

Step 1: Compare the minimal gains. Evaluate whether the minimal win in Gambles A and B differs by 1/10 or more of the maximum possible win. If this is the case, choose the gamble with the larger minimal win. Ignore the rest of the information. If the minimal gains differ by less than 1/10 of the maximum gain, move on to Step 2.
Step 2: Compare probabilities of the minimum gains. Evaluate whether the probabilities of the minimal wins differ by 1/10 or more. If this is the case, choose the gamble with the smaller of the two probabilities. If the probabilities of the minimum gains differ by less than 1/10, move on to Step 3.
Step 3: Compare the maximum gains. Evaluate which of the two gambles, A or B, has the higher maximum gain. Choose the gamble with the higher maximum gain. If the maximum gains do not differ, choose randomly.

For **losses**, the following steps have to be used:

Step 1: Compare the minimal losses. Evaluate whether the minimal losses in Gambles A and B differ by 1/10 or more of the maximum possible loss. If this is the case, choose the gamble with the smaller minimal loss. Ignore the rest of the information. If the minimal losses differ by less than 1/10 of the maximum loss, move on to Step 2.
Step 2: Compare probabilities of the minimum losses. Evaluate whether the probabilities of the minimal losses differ by 1/10 or more. If this is the case, choose the gamble with the larger of the two probabilities. If the probabilities of the minimum losses differ by less than 1/10, move on to Step 3.
Step 3: Compare the maximum losses. Evaluate which of the two gambles, A or B, has the lower maximum loss. Choose the gamble with the lower maximum loss. If the maximum losses do not differ, choose randomly.

Please note that aids like pocket calculators or paper are not allowed during the study.

Exercise

We will now walk you through a step-by-step example how to apply the three steps described above:

Gamble A	10	50%	0	50%
Gamble B	20	20%	1	80%

Step 1:	Read each of the values once.
Step 2:	Divide the largest gain by 10 and remember the result. **20/10 = 2**
Step 3:	Calculate the difference between the two smallest gains. **1 − 0 = 1**
Step 4:	If the difference between the two smaller gains (Step 3) is equal or larger than the result from Step 2, then choose the gamble with the larger smaller win; if this is not the case, continue to Step 5. **1 < 2 ➔ continue**
Step 5:	Calculate the difference between the probability of the smaller gains and remember the result. **80 − 50 = 30**
Step 6:	If the difference of the probabilities of the smaller gains (Step 5) is equal or larger than 10%, choose the gamble with the smaller chance of the smaller gain. **30% > 10% ➔ Gamble A ➔ Press A**

DOI: 10.1002/bdm

REFERENCES

Böckenholt, U., & Hynan, L. S. (1994). Caveats on a process-tracing measure and a remedy. *Journal of Behavioral Decision Making*, *7*(2), 103–117. 10.1002/bdm.3960070205.

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review*, *113*(2), 409–432. 10.1037/0033-295X.113.2.409.

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2008). Risky choice with heuristics: Reply to Birnbaum (2008), Johnson, Schulte-Mecklenbeck, and Willemsen (2008), and Rieger and Wang (2008). *Psychological Review*, *115*(1), 281–289. https://doi.org/10.1037/0033-295X.115.1.281

Brandstätter, E, & Körner, C. (2014). Attention in risky choice. *Acta Psychologica*, *152*, 166–176. https://doi.org/10.1016/j.actpsy.2014.08.008

Day, R.-F. (2010). Examining the validity of the Needleman–Wunsch algorithm in identifying decision strategy with eye-movement data. *Decision Support Systems*, *49*(4), 396–403. 10.1016/j.dss.2010.05.001.

DeAngelus, M., & Pelz, J. B. (2009). Top-down control of eye movements: Yarbus revisited. *Visual Cognition*, *17*(6-7), 790–811. 10.1080/13506280902793843.

DeWolf, M., Grounds, M. A., Bassok, M., & Holyoak, K. J. (2014). Magnitude comparison with different types of rational numbers. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(1), 71. 10.1037/a0032916.

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

Ford, J. K., Schmitt, N., Schechtman, S. L., Hults, B. M., & Doherty, M. L. (1989). Process tracing methods: Contributions, problems, and neglected research questions. *Organizational Behavior and Human Decision Processes*, *43*(1), 75–117.

Friedman, M., & Savage, L. J. (1948). The utility analysis of choices involving risk. *Journal of Political Economy*, *56*(4), 279–304.

Glaholt, M. G., Wu, M. C., & Reingold, E. M. (2010). Evidence for top-down control of eye movements during visual decision making. *Journal of Vision*, *10*(5), 15–15. 10.1167/10.5.15.

Glöckner, A., & Herbold, A.-K. (2011). An eye-tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes. *Journal of Behavioral Decision Making*, *24*(1), 71–98. 10.1002/bdm.684.

Hacking, I. (1984). *The emergence of probability: A philosophical study of early ideas about probability, induction and statistical inference*. Cambridge: Cambridge University Press.

Henson, R. (2006). Forward inference using functional neuroimaging: Dissociations versus associations. *Trends in Cognitive Sciences*, *10*(2), 64–69. 10.1016/j.tics.2005.12.005.

Huber, O., Wider, R., & Huber, O. W. (1997). Active information search and complete information presentation in naturalistic risky decision tasks. *Acta Psychologica*, *95*(1), 15–29.

Hutzler, F. (2014). Reverse inference is not a fallacy per se: Cognitive processes can be inferred from functional imaging data. *NeuroImage*, *84*(C), 1061–1069. 10.1016/j.neuroimage.2012.12.075.

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, *22*(4), 679–688. 10.1016/j.ijforecast.2006.03.001.

Johnson, E. J., & Ratcliff, R. (2014). Computational and process models of decision making in psychology and behavioral economic. In P. W. Glimcher (Ed.), *Neuroeconomics: Decision making and the brain* (2nd ed.,, pp. 35–47). Elsevier.

Johnson, E. J., Schulte-Mecklenbeck, M., & Willemsen, M. C. (2008). Process models deserve process data: Comment on Brandstätter, Gigerenzer, and Hertwig (2006). *Psychological Review*, *115*(1), 263–272. 10.1037/0033-295X.115.1.263.

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

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Publishing Group*, *13*(10), 1292–1298. 10.1038/nn.2635.

Machery, E. (2014). In Defense of Reverse Inference. *The British Journal for the Philosophy of Science*, *65*(2), 251–267. https://doi.org/10.1093/bjps/axs044

Marr, D. (1982). *Vision*. San Francisco: Freeman.

Pachur, T., Hertwig, R., Gigerenzer, G., & Brandstätter, E. (2013). Testing process predictions of models of risky choice: A quantitative model comparison approach. *Frontiers in Psychology*, *4*. 10.3389/fpsyg.2013.00646.

Payne, J. W., Braunstein, M. L., & Carroll, J. S. (1978). Exploring predecisional behavior: An alternative approach to decision research. *Organizational Behavior and Human Performance*, *22*(1), 17–44.

Pieters, R., & Wedel, M. (2007). Goal control of attention to advertising: The Yarbus implication. *Journal of Consumer Research*, *34*(2), 224–233. 10.1086/519150.

Pleskac, T. J., & Hertwig, R. (2014). Ecologically rational choice in a structured environment. *Journal of Experimental Psychology: General*, *143*(5), 2000–2019.

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, *10*(2), 59–63. 10.1016/j.tics.2005.12.004.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*, *20*, 110–122.

Reichle, E. D., Rayner, K., & Pollatsek, A. (1999). Eye movement control in reading: Accounting for initial fixation locations and refixations within the EZ Reader model. *Vision Research*, *39*(26), 4403–4411. https://doi.org/10.1016/S0042-6989(99)00152-2

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372.

Riedl, R., Brandstätter, E., & Roithmayr, F. (2008). Identifying decision strategies: A process- and outcome-based classification method. *Behavior Research Methods*, *40*(3), 795–807. https://doi.org/10.3758/BRM.40.3.795

Russo, J. E. (2011). Eye fixations as a process trace. In M. Schulte-Mecklenbeck, A. Kühberger, & R. Ranyard (Eds.), *A handbook of process tracing methods for decision research* (pp. 43–64). Oxford: Psychology Press.

Schkade, D. A., & Johnson, E. J. (1989). Cognitive processes in preference reversals. *Organizational Behavior and Human Decision Processes*, *44*, 203–231.

Schulte-Mecklenbeck, M., Kühberger, A., & Ranyard, R. (2011a). *A handbook of process tracing methods for decision research*. Taylor & Francis.

Schulte-Mecklenbeck, M., Kühberger, A., & Ranyard, R. (2011b). The role of process data in the development and testing of process models of judgment and decision making. *Judgement and Decision Making*, *6*(8), 733–739.

Schulte-Mecklenbeck, M., Sohn, M., de Bellis, E., Martin, N., & Hertwig, R. (2013). A lack of appetite for information and computation: Simple heuristics in food choice. *Appetite*, *71*, 242–251.

Su, Y., Rao, L.-L., Sun, H.-Y., Du, X.-L., Li, X., & Li, S. (2013). Is making a risky choice based on a weighting and adding process? An eye-tracking investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *6*, 1765–1780. 10.1037/a0032861.

Svenson, O. (1979). Process descriptions of decision making. *Organizational Behavior and Human Performance*, *23*(1), 86–112.

Svenson, O. (1996). Decision making and the search for fundamental psychological regularities: What can be learned from a

DOI: 10.1002/bdm

process perspective? *Organizational Behavior and Human Decision Processes*, *65*(3), 252–267.
Teuber, H. L. (1955). Physiological psychology. *Annual Review of Psychology*, 267–296.
Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323.
Velichkovsky, B. M. (1999). From levels of processing to stratification of cognition: Converging evidence from three domains of research. In B. H. Challis, & B. M. Velichkovsky (Eds.), *Stratification in cognition and consciousness* (pp. 203–235). Amsterdam: John Benjamins Publishing Company.
Wedell, D. H., & Senter, S. M. (1997). Looking and weighting in judgment and choice. *Organizational Behavior and Human Decision Processes*, *70*(1), 41–64.
Willemsen, M. C., Böckenholt, U., & Johnson, E. J. (2011). Choice by value encoding and value construction: Processes of loss aversion. *Journal of Experimental Psychology: General*, *140*(3), 303–324. 10.1037/a0023493.
Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum Press.
Zbrodoff, N. J., & Logan, G. D. (2004). What everyone finds. The problem-size effect. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 331–347). New York: Psychology Press.

Authors' biographies:

Michael Schulte-Mecklenbeck is a lecturer at the University of Bern and an associated Research Scientist at the Max Planck Institute for Human Development, Berlin. His research focuses on process tracing, food choice, heuristics, and consumer behavior.

Anton Kühberger is Professor of Psychology at the University of Salzburg, Austria. He is working at the Department of Psychology and is a member of the Centre for Cognitive Neuroscience at the University of Salzburg. He has performed extensive research in judgment and decision making, especially in the context of risk and uncertainty. His research interests also include the development of the scientific method, and meta-research.

Benjamin Gagl is a Marie-Sklodowska-Curie Fellow at the Department of Psychology at the Goethe University Frankfurt. His research focuses on eye movement methodology and neurocognitive underpinnings of reading.

Florian Hutzler is Professor of Psychology at the University of Salzburg, Austria. He is head of the Neurocognition Lab at the Centre for Cognitive Neuroscience. He has performed research in reading, dyslexia, eye tracking and fMRI.

Authors' addresses:

Michael Schulte-Mecklenbeck, Department of Business Administration, Consumer Behavior, University of Bern, Bern, Switzerland; Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany.

Anton Kühberger, Centre of Cognitive Neuroscience & Department of Psychology, University of Salzburg, Salzburg, Austria.

Benjamin Gagl, Department of Psychology & Center for Individual Development and Adaptive Education of Children at Risk (IDeA), Goethe University Frankfurt, Frankfurt, Germany.

Florian Hutzler, Centre of Cognitive Neuroscience & Department of Psychology, University of Salzburg, Salzburg, Austria.

DOI: 10.1002/bdm