

Article

Debiasing in a Minute or Less, or Too Good to Be True? The Effect of Micro-Interventions on Decision-Making Quality

Marko Kovic

Ars Cognitionis GmbH, Buckhauserstrasse 34, 8048 Zürich, Switzerland; marko.kovic@ars-cognitionis.ch;
Tel.: +41-76-335-06-17

Received: 3 April 2019; Accepted: 11 May 2019; Published: 16 May 2019



Abstract: In this study, the effects of a novel debiasing micro-intervention (an intervention that requires no training or instructions) are tested in three experiments. The intervention is epistemologically informed and consists of two questions that prompt the quantification of degrees of a belief (“How certain am I?”) and the justification of a belief (“Why?”). In all three experiments, this intervention was ineffective. Unexpectedly, however, when the micro-intervention consisted only of the justification question (“Why?”), there was a small, but noticeable positive effect in two experiments. Overall, even though the hypothesized effects were not observable, a justification prompt might be a potentially effective micro-intervention that should be explored in future research.

Keywords: cognitive heuristics; cognitive biases; debiasing; rationality; epistemic rationality

1. Introduction: The (Elusive) Quest for Debiasing

Cognitive heuristics are a well-documented trait of human cognition. Heuristics represent an efficient way for making inferences and decisions that are sufficiently precise and sufficiently accurate for maximizing utility and achieving goals [1,2]. It has been observed for decades, however, that inference- and decision-making deviates from theoretically desirable, “pure” rationality [3]. There is ample evidence that cognitive heuristics, useful though they are in general, can lead to (sometimes drastically) irrational inferences and decisions [4]. In such instances, cognitive heuristics lead to inferences and behavior that is often referred to as “biases”.

Mitigating biases through targeted “debiasing” interventions could potentially have an enormous positive impact [5]; not just for individuals, but for humankind as a whole. After all, even the most consequential decisions in politics or business, to name just two societal domains that deeply affect everyone’s wellbeing, are made by humans who are, by their very nature, bias-prone. Unfortunately, cognitive biases have proven to be more resilient than might be to our liking. One of the main challenges with effective debiasing interventions and strategies was outlined by [6]: There seems to be a correlation between the amount of (cognitive) resources required for debiasing interventions and the scope of the positive impact of debiasing interventions. Ref. [6] introduced the distinction between *procedural* and *structure modifying* debiasing interventions as a framework for distinguishing between two general types of debiasing that differ in terms of their resource requirements and the breadth of their positive impact.

On one hand, debiasing interventions that are *procedural* in nature and affect a specific decision-making process quasi-mechanistically are relatively easy to devise and implement, but they usually apply only to specific decision-making situations. This means that it can be relatively easy to change and improve a *specific* decision-making procedure, but those changes do not necessarily improve inference- and decision-making in general. Changing the procedure with which a decision is

made affects the quality of a decision primarily from “without” (the external decision-making context) rather than from “within” (a person’s cognition). Some examples debiasing interventions that are procedural in nature are checklists [7–10], mechanisms for slowing down or pausing decision-making processes [11,12], and consider-the-alternative stimuli [13,14]. Even though there is evidence that such interventions can improve decision-making in the specific decision-making contexts in which they are applied, there is no evidence that their positive effects extend to or spill over into other decision-making contexts. For example, there is evidence that consider-the-alternative stimuli can help avoid the confirmation bias [15] and related biases in abductive [16] decision-making situations such as clinical diagnostic reasoning [17,18] or intelligence assessment [19], but there is no plausible evidence to suggest that such targeted interventions improve inferences and decision-making beyond the narrow contexts they are applied to.

On the other hand, debiasing interventions that are *structure modifying* in nature and that allow individuals to modify the internal representation of the decision-making situation have a broader, more generalized scope, but they tend to be more demanding in terms of resources such as time and active cognitive effort. Debiasing interventions that are structure modifying in nature aim to affect inferences and decision-making from “within”: Rather than changing the external inference- and decision-making procedure, structure modifying debiasing interventions aim to, generally speaking, improve cognition. Such debiasing interventions aim to tackle the metaphorical root of the problem, cognitive biases themselves, rather than change external decision-making factors in order to reduce the probability of biased decision-making.

Structure modifying debiasing interventions have been described as being aimed at *metacognition* [20,21], the knowledge or cognition about cognitive phenomena [22]. If we understand structure modifying debiasing interventions as interventions aimed at metacognition, then their goal can be understood as the general triggering of metacognitive experiences. Such experiences should, ideally, improve the quality of beliefs as well as the quality of decision-making. There is some evidence that metacognitive experiences, such as the ease of retrieval of relevant information, can affect the prevalence of cognitive biases [23,24]. More generally, the theoretical perspective of dual process theory [25,26] lends credibility to this metacognitive view on structure modifying debiasing interventions. If there is (an idealized) mode of slower and more deliberate thinking that is less reliant on heuristics, then structure modifying debiasing interventions aimed at inducing metacognitive experiences should be beneficial for entering that cognitive mode.

A considerable part of the existing literature on structure modifying, metacognition-related debiasing interventions is making plausible suggestions for debiasing interventions, but those suggestions are often not directly being put to the empirical test [17,27]. There is, however, a growing body of recent empirical explorations of carefully designed interventions that suggest that this type of more demanding and more general metacognition-related debiasing intervention can be effective [28–31]. However, there are also studies that have found no positive effects [32,33].

One of the main drawbacks of structure modifying debiasing interventions is that they are more costly to implement in terms of time and cognitive effort. The term “cognitive forcing” that has been proposed as a descriptor for structure modifying debiasing interventions [34] captures that cost in a figurative way: Structure modifying debiasing interventions usually require decision-makers to learn new information and to subsequently “force” themselves to apply that knowledge in decision-making situations in which they would not normally do so.

Procedural and structure modifying debiasing interventions both have benefits and drawbacks. Given their respective benefits and drawbacks, a potential “third way” for debiasing intervention seems potentially attractive: Interventions that are easy to implement and apply, but that are not limited to specific and narrow decision-making contexts.

A “Third Way” for Debiasing Interventions?

Procedural debiasing interventions are easy to devise and implement, but their effectiveness is limited. The effectiveness of structure modifying debiasing interventions that target metacognition is probably broader and more generalized, but they are more costly to devise, implement, and apply.

In theory, debiasing interventions that combine some of the benefits of procedural and structure modifying interventions while avoiding some of their disadvantages could be an attractive “third way” for debiasing interventions. A “third way” debiasing intervention would need to be simple and easily deployable in practice while still targeting metacognitive experiences. In a recent study, Ref. [35] tested an intervention that has such “third way” properties. The researchers designed a relatively simple checklist that can be applied in a procedural manner, without the need for time-consuming training and instructions. However, the checklist was designed to induce a form of metacognitive experience by triggering clinicians to actively think about the way they are making inferences and decisions in a given decision-making situation.

Even though specific intervention was designed specifically for the professional context of clinicians, the “third way” of debiasing interventions, I believe, merits further attention. Debiasing interventions that follow this “third way” approach could potentially prove to be a useful general addition to the repertoire of debiasing interventions. For example, “third way” debiasing interventions could be a viable option in real-world decision-making contexts that do not allow for costly structure modifying interventions, but that still require a form

The goal of the present study is to introduce and test a novel “third way” debiasing intervention. By design, the intervention of interest is compact and almost trivially easy to implement. For that reason, I refer to it as a “micro-intervention”. The micro-intervention that I introduce and test is not merely a compact tool that offers information about cognitive biases (Not least because, in one case, such a “third way” intervention that consisted of a checklist with information on cognitive biases failed to reduce error [36]), but rather an intervention that aims to increase *epistemic rationality* with little to no effort.

2. An Epistemologically Informed Micro-Intervention

2.1. The Importance of Epistemic Rationality

Cognitive biases are an impediment to rationality, and debiasing aims to increase rationality. However, what, exactly, do we mean when we refer to “rationality”? A preliminary, *ex negativo* answer would be to define rationality as the state of human cognition and action given the absence of cognitive biases. Such a definition, however, is not only lacking, but it is fallacious (This would be a case of denying the antecedent. The absence of cognitive biases does not automatically result in rationality because there are other factors that can cause irrational beliefs and actions.). An imaginary perfectly rational actor might not be prone to cognitive biases, but there is more to rationality than just the absence of biases.

The exact meaning and definition of rationality is not a settled question. In contemporary academic discourse, two types of rationality are generally of interest: *Instrumental rationality* and *epistemic rationality*. Instrumental rationality means that an actor makes decisions in such a way that they achieve their goals [37]. Instrumental rationality is the foundation of rational choice theory, and research on cognitive heuristics and biases is usually concerned with the discrepancy between theoretically expected rational choice decision-making and the observable real-world decision-making [38]. However, instrumental, goal-oriented decision-making is not all there is to rationality. When we are making conscious and deliberate decisions, for example, those decisions are based on a layer of doxastic (belief-like) attitudes towards reality, most important of which are, arguably, beliefs themselves. The concern with the quality of our beliefs falls into the domain of epistemic rationality.

Epistemic rationality means that an actor is concerned with adopting beliefs about reality that are as accurate and as precise as possible [39]. That does not, however, mean that an imaginary epistemically rational actor is epistemically rational simply because they happen to hold true beliefs. For example, a belief can “accidentally” be true, but the reasons for holding that belief might be quite irrational (Such constellations can generally be described as “epistemic luck” [40]). I might happen to believe that the Earth is spherical (a true belief), but my justification for that belief might be that all celestial bodies are giant tennis balls (an irrational justification). Epistemic rationality, therefore, is concerned with the process of arriving at a belief and not necessarily (only) with the objective veracity of a belief.

Cognitive biases affect both instrumental and epistemic rationality. Biases can prevent us from achieving goals, and they can reduce the quality of our belief-formation. From a cognitive perspective, the root cause of dysfunctional instrumental rationality are problems with epistemic rationality; irrational decisions are usually a result of inappropriate doxastic attitudes such as beliefs. “Third way” debiasing interventions that aim to induce beneficial metacognitive experiences should therefore attempt to increase epistemic rationality and not just mitigate deficits of instrumental rationality. It should be noted, for the sake of completeness, that the distinction between instrumental and epistemic rationality might be somewhat moot altogether. Epistemic rationality is the underpinning of instrumental rationality, but it can be argued that epistemic rationality itself is instrumental in nature. After all, by being epistemically rational, an actor pursues the goal of arriving at accurate and precise beliefs [41]. For the purpose of this study, I uphold the distinction between instrumental and epistemic rationality in order to make it clear that the goal of the micro-intervention that is tested is to increase the quality of belief-formation. Ultimately, however, increasing the quality of belief-formation can be regarded as an instrumental goal.

2.2. An Epistemically Rational Actor in Practice

A debiasing intervention aimed at reducing the susceptibility to cognitive biases by increasing epistemic rationality should address two specific dimensions of epistemic rationality: *Justification* and *credence*.

Justification means that an actor is rational in holding a belief if the actor has reasons for doing so, for example because there is empirical evidence that points to the veracity of the belief [42,43]. Justificationism in general and evidentialism in particular are not universally accepted in contemporary epistemology [44], but they represent a fairly uncontroversial “mainstream” view. In practical terms, the dimension of justification means that an actor is rational when they have evidence in support of the belief they are holding. For example, person A and person B might both believe that the Earth is spherical, but person A might simply rely on their intuition, whereas person B might cite some astronomical and geographical evidence as a justification. In this case, even though A and B both happen to hold a true belief, we would regard person B as more rational because their reasons for holding the belief they are holding are better (An important ongoing debate in epistemic justificationism is the question of whether rational justification is primarily internal or also external to a person’s cognition [45]. I tend to regard the internalist view proposed by [46] as convincing, since the question of rational belief-formation is, in many demonstrable cases, divorced from the question of whether the beliefs a person arrives at are, in fact, true. However, the internalism vs. externalism debate is not central to the concept of “third way” debiasing interventions that are explored in this paper.)

Credence refers to the idea that an actor should not merely have belief or disbelief in some proposition. Instead, it is more rational for an actor to express their attitude towards a proposition in terms of *degrees of belief*, depending on the strength of the available evidence [47]. In principle, degrees of belief could be expressed on any arbitrary scale, but the most suitable scale for expressing beliefs is probability. Indeed, one prominent school of thought regards probabilities as ultimately nothing more than expressions of degrees of belief [48]. This view of probability is the foundation of Bayesian flavors of epistemology [49,50]. In practical terms, the dimension of credence means that an epistemically

rational actor should probabilistically express how strongly they believe what they believe, given the available evidence. For example, if person A from the example above expressed that they believed the Earth to be round with probability 1, we might suspect that person A is being severely overconfident, given the quality of their justification (merely intuition).

2.3. Designing a “Worst-Case” Micro-Intervention

In the preceding subsection, I have outlined the argument that a debiasing intervention aimed at increasing epistemic rationality should target the dimension of justification and credence. There are many ways in which that conceptual basis of an intervention could be put into practice. For example, an intervention could consist of a formal apparatus whereby people have to explicitly state reasons for a belief they hold as well as explicitly quantify their belief on a scale from 0 to 1. Such an apparatus might represent an interesting (and attention-worthy) intervention in its own right, but it would hardly qualify as a *micro-intervention*. A veritable micro-intervention has to, I believe, be so trivially easy to implement in a decision-making situation that it requires zero preparation and essentially zero effort. We can describe such a design principle as a “worst-case” intervention: An intervention that can be used in real-world decision making situations even when there are no resources to implement interventions and even when the motivation on the part of the decision-makers themselves is low.

For the purpose of the present study, I design the worst-case micro-intervention as a prompt that consists of merely two questions:

- How certain am I?
- Why?

The goal of this micro-intervention is to trigger metacognitive experiences that pertain to the dimensions of credence (How certain am I?) and justification (Why?). If this micro-intervention is effective, it should produce a noticeable improvement of decision-making even in worst-case circumstances (low motivation on the part of the decision-makers paired with no training or formal instructions).

In an experimental setting, the effectiveness of this micro-intervention can be compared against a control group without the micro-intervention. For the sake of completeness, another micro-intervention should be added to the experimental setting: a micro-intervention that consists of a mere *stopping intervention*. It is possible that the main micro-intervention that I test is effective, but that the true effect is an underlying stopping mechanism and not the epistemological aspects of the micro-intervention. Furthermore, the micro-intervention of interest consists of two separate questions. For the sake of completeness, it is appropriate to test variants of micro-interventions that consist of only one of the two questions.

2.4. The Potential Real-World Benefits of the Proposed Micro-Intervention

The “third way” micro-intervention described above is tested under experimental “laboratory” conditions; the design and results are presented below. However, first, a brief discussion of the potential real-world benefits of such an intervention is in order. If we bother to evaluate the evidence this intervention (or that of “third way” interventions in general), we should be motivated by potential tangible benefits for real-world decision-making. What might those benefits be?

It is unrealistic to expect the proposed micro-intervention to have overall large positive effects. From a purely practical perspective, we can expect problems such as attrition in the usage of the micro-intervention if it is introduced as a veritable micro-intervention (without extensive training or instruction, and without controlling whether it is applied). For example, if an organization were to introduce the proposed micro-intervention as a tool to all employees with management status (for example, in the form of an email memo or an agenda item during a meeting), many of those employees might forget about the micro-intervention after a while, or they might even fail to see its benefit in the first place. On a more conceptual level, we can expect that not all bias-related rationality

deficits can be mitigated with the proposed micro-intervention because the strength of specific biases might be orders of magnitude greater than the metacognitive experiences that are induced by the micro-intervention.

I believe that the real-world benefit of the proposed micro-intervention would lie in a small relative reduction of the susceptibility to biases through what I would label as *habitualized irritation*. It is unrealistic to expect the proposed micro-intervention to produce great ad hoc improvements in epistemic rationality, but prompting oneself into thinking about one's confidence in and one's justification for a judgement or decision might shift some decisions away from a gut-level and more towards a deliberate cognitive effort.

The real-world benefit of the proposed micro-intervention (and of "third way" interventions in general) is probably greatest in organizational contexts in which the intervention is integrated into decision-making processes, whereby decision-makers are regularly prompted to use the intervention as a tool for their own decision-making. In such a setting, the risk of biased decision-making might still only slightly decrease per singular decision, but from a cumulative and aggregate perspective over time, the beneficial effects might be nontrivial.

3. Hypotheses

The effectiveness of the proposed micro-intervention can be tested empirically in an experimental setting by comparing the decision-making quality of an intervention group with that of a control group without any intervention. However, for the sake of completeness, additional comparisons can be made. Given the theoretical background, the micro-intervention that consists of both questions should have a greater effect than just one question on its own (either "How certain am I?" or "Why?"). However, even only one of the two questions should have a positive impact compared to the control group without an intervention. In addition, it is conceivable that the micro-intervention or the partial micro-interventions (that contain only one of the two questions) do have an effect, but that the effect is merely the result of slowing down a bit. Therefore, an additional stopping intervention should be tested against the micro-intervention, the partial micro-intervention, and the control group.

These assumptions lead to the following three hypotheses:

Hypothesis 1 (H1). *The combined probability + justification intervention is the most effective intervention.*

Hypothesis 2 (H2). *The mere stopping intervention is less effective than the combined probability + justification intervention.*

Hypothesis 3 (H3). *All interventions (mere stopping, probability, justification, probability + justification) will have a positive effect compared to the control group without any interventions.*

At first glance, Hypothesis 2 (H2) might seem redundant given Hypothesis 1 (H1), but its purpose is to explicitly describe the expectation that

$$\text{probability+justification} > \text{stopping}$$

and not

$$\text{probability+justification} \geq \text{stopping}.$$

4. Design, Data, Methods

4.1. Design

In order to test the hypotheses, I have conducted three separate experiments. Each experiment was designed to test the effects of the micro-interventions on a different cognitive bias. Experiment 1 was designed around the *gambler's fallacy* [51,52], experiment 2 around the *base rate fallacy* [52],

and experiment 3 around the *conjunction fallacy* [53]. I have chosen these three probability-oriented biases because they allow for the creation of clear-cut, yes-or-no scenarios and questions.

Experiment 1 consisted of the following scenario:

John is flipping a coin (It's a regular, fair coin.). He has flipped the coin four times - and each time, heads came up. John is about to flip the coin a fifth time.

What do you think is about to happen on the fifth coin toss?
(*The intervention was placed here*)

- Tails is more likely to come up than heads.
- Tails is just as likely to come up as heads.

The correct answer for experiment one is that tails is just as likely to come up as heads. The gambler's fallacy in this case would be to answer that Tails is more likely to come up (The probability of tails is unaffected by prior outcomes.). The order of the two answer options was randomized. The interventions that were placed after the question are summarized in Table 1.

Table 1. Interventions used in the three experiments.

Condition	Intervention
Control	-
Mere stopping	<i>Before you pick your answer, take a few moments to think.</i>
Probability	<i>Before you pick your answer, ask yourself: How certain am I that this is the right answer?</i>
Justification	<i>Before you pick your answer, ask yourself: Why do I think this is the right answer?</i>
Probability+ justification	<i>Before you pick your answer, ask yourself: How certain am I that this is the right answer? Why do I think this is the right answer?</i>

Each intervention represents one experimental conditions. In total, there were five experimental conditions to which the participants were randomly assigned.

Experiment 2 consisted of the following scenario:

The local university has started using a computer program that detects plagiarism in student essays. The software has a false positive rate of 5%: It mistakes 5% of original, non-plagiarized essays for plagiarized ones. However, the software never fails to detect a truly plagiarized essay: All essays that are really plagiarized are detected - cheaters are always caught.

Plagiarism at the university is not very common. On average, one out of a thousand students will commit plagiarism in their essays. The software has detected that an essay recently handed in by John, a university freshman, is plagiarized.

What do you think is more likely?
(*The intervention was placed here*)

- John has indeed committed plagiarism.
- John is innocent and the result is a false positive.

The correct answer for experiment 2 is that it is more likely that John is innocent and the result is a false positive. Given the information in the scenario, the actual probability that John is indeed guilty

is only 0.02. The information provided in the scenario of experiment 1 is adapted from a classical test of the base-rate fallacy [54,55]. The correct probability can be calculated by applying Bayes' rule

$$\Pr(A | B) = \frac{\Pr(B | A) \times \Pr(A)}{\Pr(B)}$$

as

$$\Pr(\text{plag} | \text{pos}) = \frac{\Pr(\text{pos} | \text{plag}) \times \Pr(\text{plag})}{\Pr(\text{pos} | \text{plag}) \times \Pr(\text{plag}) + \Pr(\text{pos} | \neg\text{plag}) \times \Pr(\neg\text{plag})}$$

where *plag* stands for plagiarism and *pos* for positive test result by the plagiarism detection software. When the numbers provided in the scenario are plugged into the equation, we arrive at the probability of around 0.02, or around 2%:

$$\Pr(\text{plag} | \text{pos}) = \frac{1 \times 0.001}{1 \times 0.001 + 0.05 \times 0.999} = 0.0196.$$

The order of the answer options in experiment 2 was randomized. The interventions (and therefore the conditions) for experiment 2 were the same as for experiment 1, as summarized in Table 1. Experiment 3 consisted of the following scenario:

Lisa started playing tennis when she was seven years old. She has competed in several amateur tennis tournaments, and she even won twice.

Which of the following descriptions do you think is more likely to be true for Lisa?

(The intervention was placed here)

- Lisa is a lawyer.
- Lisa is a lawyer. She play tennis on the weekends.

The correct answer for experiment 3 is that "Lisa is a lawyer" is more likely to be true. This experiment is a slight adaptation of the famous Linda experiment [53]. The order of the answer options in experiment 3 was randomized, and the interventions (and therefore the conditions) for experiment 3 are, once again, the same as in experiments 1 and 2, as summarized in Table 1.

4.2. Participants

Participants for the three experiments were recruited via the online crowdworking platform Clickworker [56,57]. For each experiment, the participation of 500 residents of the United States was commissioned on the crowdworking platform. Each participant was remunerated with 0.15 Euro. The mean completion time for experiment 1 was 54.4 s, for experiment 2 was 72.5 s, and for experiment 3 was 57.2 s.

In total, there were 508 participants in experiment 1, 501 participants in experiment 2, and 500 participants in experiment 3. The participants were similar in terms of age (means of 34.2, 33.8, 33.8 and standard deviations of 12.4, 12.6, 12.3 years for experiments 1, 2, and 3) and sex ratio (62.6%, 63.1%, 62.8% female for experiments 1, 2, and 3) in all three experiments. The experiments were conducted with the survey software LimeSurvey [58]. In each experiment, participants were randomly assigned to one of the five experimental condition. The number of participants per experimental condition in each experiment is summarized in Table 2.

Table 2. Number of participants per experimental condition in each experiment.

	Experiment 1	Experiment 2	Experiment 3
Control	92	100	98
Mere stopping	93	108	99
Probability	106	96	106
Justification	111	99	99
Probability + justification	106	98	98

The data was collected between 8 August and 9 August 2018. The data was not manipulated after the collection; all data that was collected is part of the data analysis.

4.3. Main Data Analysis

The experiments represent a between-subjects design with five levels. Since the participants in each experiment could either choose the correct or the incorrect answer, the observed data can be thought of as being generated by an underlying binomial distribution. The probability for a “success” (correct answer) is the unobserved variable of interest. I estimate this parameter as a simple Bayesian model in the following manner:

$$x \sim \text{binomial}(n, \theta),$$

$$\theta \sim \beta(1, 1).$$

In this model, x stands for the number of successes, n for the number of trials, and θ for the probability for success. The prior $\beta(1, 1)$ is a flat prior, declaring that any value between 0 and 1 is equally probable for θ . In practical terms, a flat prior means that the posterior distribution is determined entirely by the observed data.

I estimate θ for every experimental condition in the three experiments, resulting in a total of 15 model estimates (five per experiment). The resulting posterior distributions show whether and how θ differs between experimental conditions. This approach is not a statistical test but a parameter estimate that quantifies uncertainty given the data.

The model estimations were performed with the probabilistic language Stan [59] called from within the statistical environment R [60]. For each model estimate, four MCMC (Markov chain Monte Carlo) chains were sampled, with 1000 warmup and 1000 sampling iterations per chain. All model estimates converged well, as indicated by potential scale reduction factors [61] of $\hat{R} = 1$ for all parameters in all models. The model estimates are reported in the following results section as graphical representations of the full posterior distributions and as numeric summaries in tabular form.

4.4. Complementary Data Analysis

For the complementary data analysis, the results of the experiments are estimated in the form of the following Bayesian logistic regression:

$$y \sim \text{Bernoulli}(\mu),$$

$$\mu = \text{logit}(\alpha + \beta x),$$

$$\alpha \sim t(3, 0, 2.5),$$

$$\beta \sim t(3, 0, 2.5).$$

The benefit of estimating this regression model is that all parameter estimates (per experiment) are contained within a single model (They are not estimated separately, as in the main data analytic procedure.). The resulting model estimates allow for a more direct comparison of the effects of the

different interventions. In addition, as a step towards methodological triangulation, discrepancies between the primary and the complementary data analysis yields a more precise picture.

In the logistic regression models, the priors for the parameter estimates are vague, non-committal priors that give large weight to the data [62,63]. For the noise distribution, I use Student's *t*-distribution rather than a normal distribution because the former is more robust against outliers [64].

The three model estimations (one per experiment) are performed in the same manner as the model estimation of the primary data analysis: For each model, four MCMC chains were sampled, with 1000 warmup and 1000 sampling iterations per chain. All estimated parameters in all three models have potential scale reduction factors of $\hat{R} = 1$, indicating that the models converged well. The model estimation were again performed with the probabilistic language Stan from within the the statistical environment R, but in this instance, the analysis in Stan was executed through the package BRMS within R [65] (Packages such as BRMS do not change how Stan works. They are interfaces that make the model and data specification in Stan easier.).

4.5. Pre-Registration

The hypotheses, the research design, and the data analysis procedure were pre-registered with the Open Science Framework prior to the collection of the data [66]. The complementary data analysis presented in Appendix A has not been pre-registered.

5. Results

5.1. Experiment 1: Gambler's Fallacy

The posterior distributions for experiment 1 are depicted in Figure 1.

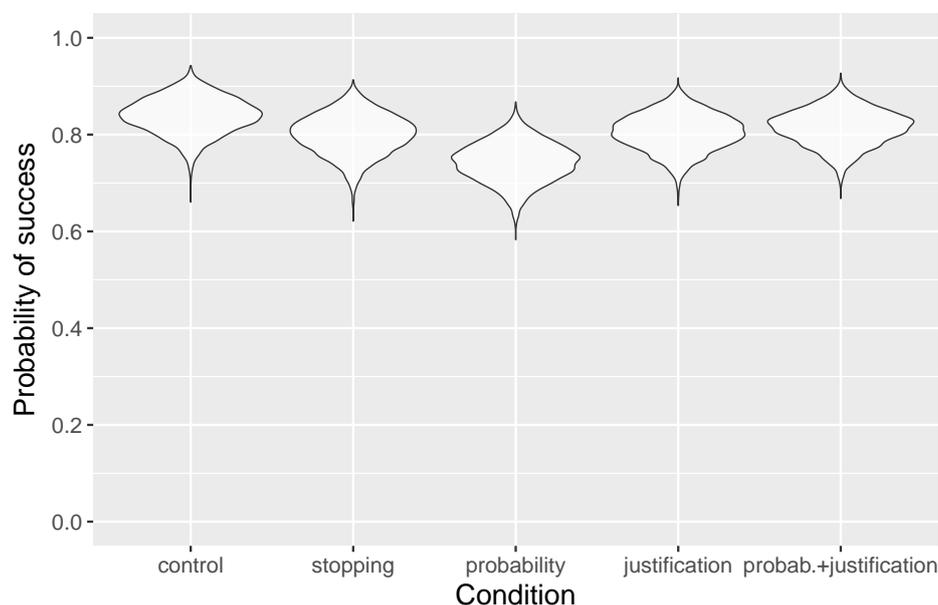


Figure 1. Posterior distributions for experiment 1.

Three pieces of information are visible in Figure 1. First, all experimental conditions have a relatively high probability of success, at around 0.8. Second, the control condition has the highest average probability. Third, most interventions had a slight *negative* effect on the probability of success—the micro-interventions did not decrease, but *increase* the probability of an error in the decision-making.

There is, however, considerable overlap among the posterior distributions. The more overlap there is, the more improbable it is that there is any meaningful difference between experimental conditions. The numeric summary in Table 3 shows that, if the criterion of 95% credible intervals [67] is employed, the only potentially meaningful difference is that between the control and probability conditions.

Table 3. Numeric summary of the model estimates for experiment 1.

	Mean	2.5% Quantile	50% Quantile	97.5% Quantile
Control	0.84	0.76	0.84	0.91
Mere stopping	0.80	0.71	0.80	0.87
Probability	0.74	0.66	0.74	0.82
Justification	0.80	0.66	0.81	0.87
Probability + justification	0.81	0.74	0.82	0.88

5.2. Experiment 2: Base-Rate Neglect

The posterior distributions for experiment 2 are depicted in Figure 2.

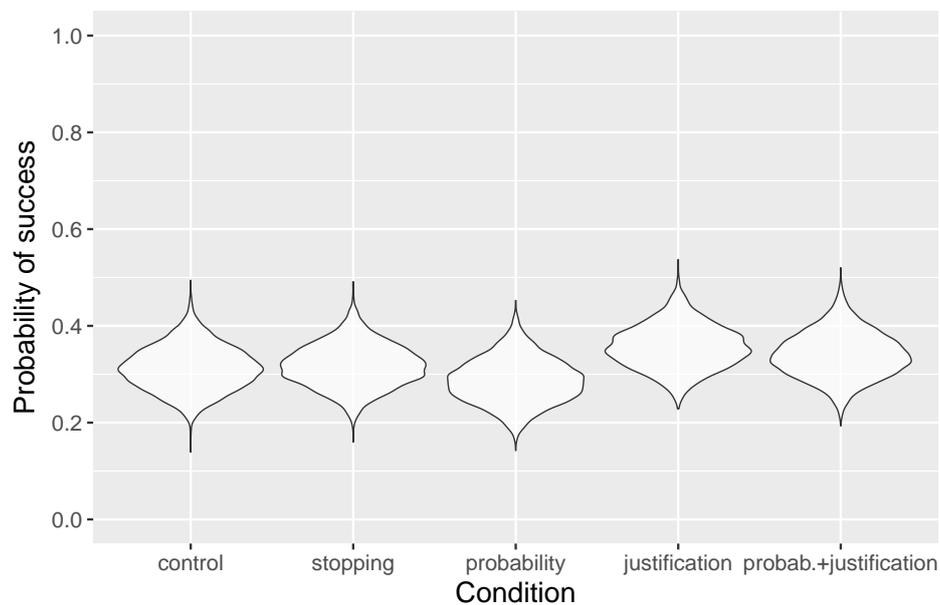


Figure 2. Posterior distributions for experiment 2.

Figure 2 offers several pieces of information. First, the average probability of success across all conditions is much lower than in experiment 1, in the area of around 0.3. Second, the probability intervention has once again resulted in the worst outcome across the five conditions. Third, however, the justification condition has resulted in a potential slight improvement compared to the control condition.

There is, however, once again considerable overlap between the posterior distributions of the five conditions. The summary in Table 4 shows that the potential positive effect of the justification intervention is unlikely to be meaningful, since there is considerable overlap between the posterior distributions even if the 95% credible interval criterion is applied.

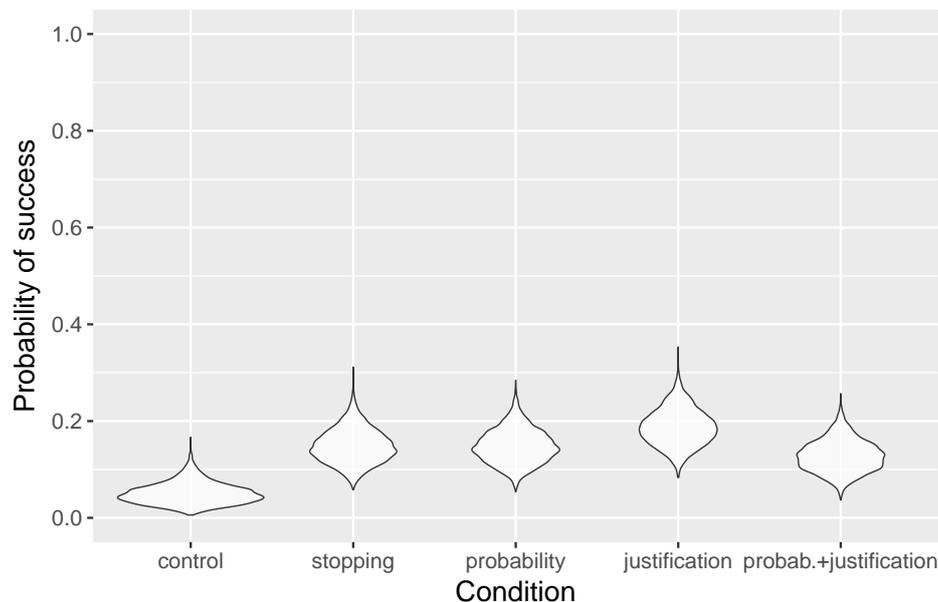
Table 4. Numeric summary of the model estimates for experiment 2.

	Mean	2.5% Quantile	50% Quantile	97.5% Quantile
Control	0.31	0.23	0.31	0.41
Mere stopping	0.32	0.23	0.32	0.41
Probability	0.29	0.20	0.29	0.38
Justification	0.36	0.27	0.35	0.45
Probability + justification	0.34	0.25	0.34	0.44

5.3. Experiment 3: Conjunction Fallacy

The posterior distributions for experiment 3 are depicted in Figure 3.

Several pieces of information are visible in Figure 3. First, the average probability of success across conditions is the lowest in experiment 3. Second, all interventions resulted in some positive effects compared to the control condition, but there is overlap between the control condition and the interventions. Third, the biggest positive effect was produced by the justification intervention, as was the case in experiment 2.

**Figure 3.** Posterior distributions for experiment 3.

Compared with experiments 1 and 2, the overlap between the posterior distributions of the control condition and the interventions in experiment 3 is less pronounced. The summary in Table 5 shows that, if the 95% credible interval criterion is applied, the justification as well as the probability + justification intervention have resulted in meaningful positive effects compared to the control condition (However, the differences between the effects of the justification and the probability + justification interventions themselves are not as pronounced.).

Table 5. Numeric summary of the model estimates for experiment 3.

	Mean	2.5% Quantile	50% Quantile	97.5% Quantile
Control	0.05	0.02	0.05	0.1
Mere stopping	0.15	0.09	0.15	0.22
Probability	0.15	0.09	0.15	0.22
Justification	0.19	0.12	0.19	0.27
Probability + justification	0.13	0.07	0.13	0.20

5.4. Complementary Data Analysis

The results of the complementary data analysis are presented in Appendix A. The reference group in each model is the control group, and the estimates for the four intervention conditions represent changes in the log-odds of choosing the correct answer.

As can be seen from a visual comparison of Figure 1 with Figure A1, of Figure 2 with Figure A2, and of Figure 3 with Figure A3, the results essentially paint the same picture. There is no discrepancy between the main and the complementary data analysis.

6. Discussion

6.1. Too Good to Be True

The title of this paper asks the question whether a micro-intervention can debias in a minute or less, or whether such a proposition is too good to be true. Given the experimental results, the answer seems clear: A trivially simple micro-intervention that still has a non-trivial effect is too good to be true.

In experiment 1, the micro-intervention of interest, the combined probability + justification intervention, actually caused harm by slightly decreasing the probability of choosing the correct answer. In experiments 2 and 3, the combined probability + justification intervention did not cause harm, but it was also not the most effective intervention. The isolated justification intervention worked best in experiments 2 and 3: Asking “Why?” worked noticeably better than asking “How certain am I?” and “Why?” at the same time. Overall, these results mean that all three hypotheses are clearly rejected.

6.2. What Can Be Salvaged?

The result of experiments 2 and 3 was surprising, given the theoretical background. The fact that the micro-intervention “Why?” seems to have worked best in these two experiments could indicate that triggering a metacognitive experience surrounding justification works better than triggering probabilistic reasoning. Justifying a belief might be a more natural, intuitive process than quantifying the degree of belief.

It would not be altogether surprising if probabilistic reasoning as a metacognitive experience was indeed difficult to induce ad hoc. After all, numerous cognitive biases are related to errors in probabilistic reasoning, among them the fallacies that were used in the three experiments of this study (the gambler’s fallacy, the base-rate fallacy, and the conjunction fallacy). If it were that easy to engage in probabilistic reasoning, those probability-related cognitive biases might not be as prevalent as research suggests they are. Expressing confidence in or the strength of one’s beliefs in probabilistic terms might not be suitable for micro-intervention debiasing formats because that kind of probabilistic reasoning might require some prior instructions and training.

6.3. Limitations of the Present Research Design

The research design of this study has several limitations and flaws. In experiment 1, all interventions had a negative effect; the control group did best. The interventions therefore seem to have caused a what can be described as a backfire effect. Backfire effects of debiasing interventions have been observed before [68]. However, in view of the fact that the overall prevalence of correct

answers across all groups was rather high, there might be a problem with the experiment itself. The experimental scenario might have been too poorly phrased, or the participants might have been familiar with the gambler's fallacy through prior exposure to similar experiments.

More generally, the research design of this study is a "worst-case" design. Participants in the different experimental conditions were presented a prompt (the respective intervention), but in hindsight, it remains unknown whether the participants actually followed the prompt. That is an intentional design choice because including formal measurements for the prompts would have transformed the micro-intervention into a more complex intervention. However, that design choice has its downsides. The participants in this study were incentivized to simply click through to the end of the experiment in order to collect their prize. They had no "skin in the game": whether or not they followed the prompts had no impact on their reward—the participants had no incentive to engage with the experiments in a serious and attentive manner because the outcome was the same even if they mindlessly clicked through the experiment. Given that the material reward is generally the primary motivation for crowdworking [69], such inattentive participation seems likely.

Of course, since the experiments in this study are randomized, it can be expected that these design downsides were uniformly distributed across the five experimental groups. However, the "worst-case" design setting of this study probably means that there is more noise in the data than there would have been if, for example, the interventions in the form of prompts were less subtle and the participants had to, for example, at least acknowledge having read the prompts (by clicking "OK", for example).

Another limitation of the research design is the choice of biases for the three experiments. I opted for probability-related biases mainly for the sake of convenience (Probability-related biases with clear correct answers can be easily implemented for experimental settings.). However, as has been long argued, different kinds of biases might require different kinds of debiasing interventions [70]. In theory, any positive effect of the tested micro-intervention should extend to other types of biases, since the "third way" logic of the intervention is one of universal applicability (The interventions aims to increase epistemic rationality in general rather than reduce the risk of specific biases.). However, the present research design with its "monoculture" of probability-related biases does not allow for inferring anything about the effects when other types of biases are at play.

6.4. The Present Results and the Overall Debiasing Literature

The contribution of the present study to the debiasing literature is twofold. In *empirical* terms, it adds little. In *conceptual* terms, however, it might add some perspective.

All hypotheses that I set out to test were clearly refuted by the experimental evidence, and the main positive effect across all interventions—the mere justification intervention showed a potential improvement in two out of three experiments—is modest. These empirical findings fit with the overall picture that the debiasing literature paints: Debiasing is probably not a hopeless effort, but there is no debiasing silver bullet that can easily mitigate the risks of cognitive biases.

The conceptual perspective might be the greater contribution of this study. The debiasing literature is mostly fragmented, without noticeable strong conceptual underpinnings. In a sense, of course, that is a good thing: Empirical trial-and-error is useful and necessary, given that there is a plethora of cognitive biases that can decrease decision-making quality in many real-world decision-making contexts. At the same time, however, greater conceptual clarity could help prioritize and guide future research efforts. In this study, I have applied the distinction between procedural and structure modifying interventions as proposed by [6], and I have added the category of "third way" interventions that aim to have the generalizability of structure modifying interventions while maintaining the simplicity of procedural interventions. More importantly, I regard structure modifying as well as "third way" interventions as being concerned mainly with increasing epistemic rationality, whereas procedural interventions are concerned with increasing instrumental rationality. Nearly all existing studies on debiasing (implicitly) focus on instrumental rationality and rational choice theory. A greater appreciation for the conceptual

distinction between instrumental and epistemic rationality could add more structure to future research, and it could be a conceptual foundation for the development of novel debiasing strategies.

6.5. Directions for Future Research

Even though the empirical results of this study paint a sobering picture, future research on “third way” debiasing interventions is still in order. Research on “third way” debiasing interventions is still relatively scarce, and the failure of the specific intervention tested in this study should not be interpreted as a failure of “third way” debiasing in general.

Perhaps the major lesson of the present study is that trivially simple micro-interventions that require zero learning and instructions but that still have meaningful positive impact might be little more than wishful thinking. “Third way” interventions probably need to have some non-trivial level of cognitive or other cost in order to have the potential for meaningfully inducing beneficial metacognitive experiences. That does not mean that we have to abandon the idea of small “third way” interventions altogether, but small debiasing interventions might be more effective when they are not quite as “micro” as the one tested in this study. For example, the intervention(s) tested in this study could be combined with a brief instruction on how to use the tool at hand, and why that is a good thing to do. That way, the intervention might be more salient, and potential effects could be more easily detectable.

The failure of the credence part of the micro-intervention that was tested represents another area that merits further attention. Quantifying beliefs seems not to be something that comes naturally or intuitively, making an instruction-less debiasing intervention that contains such a prompt ineffective or, as in this study, counterproductive. However, that does not mean that people have no useful probabilistic intuitions at all. For example, common verbal expressions for probability and uncertainty roughly, but plausibly map onto the numeric probability spectrum of 0 to 1 [71]. Expressing credence might therefore be more successful if it is devised in a way that taps into that latent probabilistic intuition.

Another important general question to explore in future research is whether and how epistemologically informed debiasing interventions can be devised and implemented. Epistemologically informed interventions are interventions that do not primarily aim to mitigate rationality deficits in the form of susceptibility to (some) cognitive biases, because their goal is to *increase* epistemic rationality in general. That increase in epistemic rationality, in turn, should result in a reduced susceptibility to cognitive biases. In other words: Interventions that increase epistemic rationality are not tools that help avoiding specific reasons for making specific mistakes, but tools that should universally increase the capacity for arriving at justified beliefs. Future research should explore how such debiasing interventions can be designed, and how effective they are at achieving the goal of debiasing. Doing so would result in greater theoretical diversity. The debiasing literature is currently (implicitly) mostly concerned with instrumental rationality and rational choice theory, but the epistemological perspective quality of belief-formation is just as important and scientifically deserving.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A. Results of the Complementary Data Analysis

Appendix A.1. Logistic Regression for Experiment 1

The estimation results of the logistic regression model for experiment 1 are summarized in Figure A1.

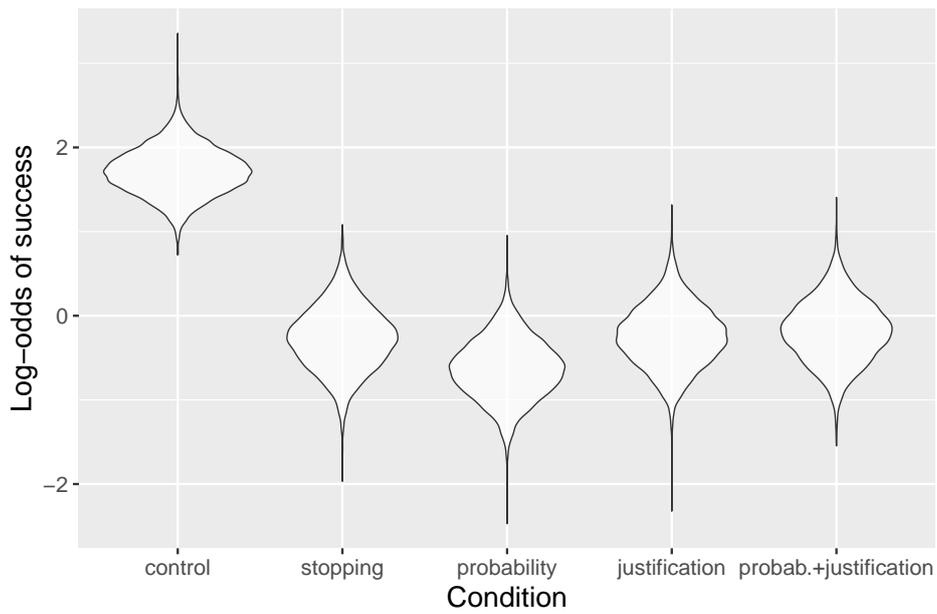


Figure A1. Posterior distributions for the logistic regression of experiment 1.

Appendix A.2. Logistic Regression for Experiment 2

The estimation results of the logistic regression model for experiment 1 are summarized in Figure A2.

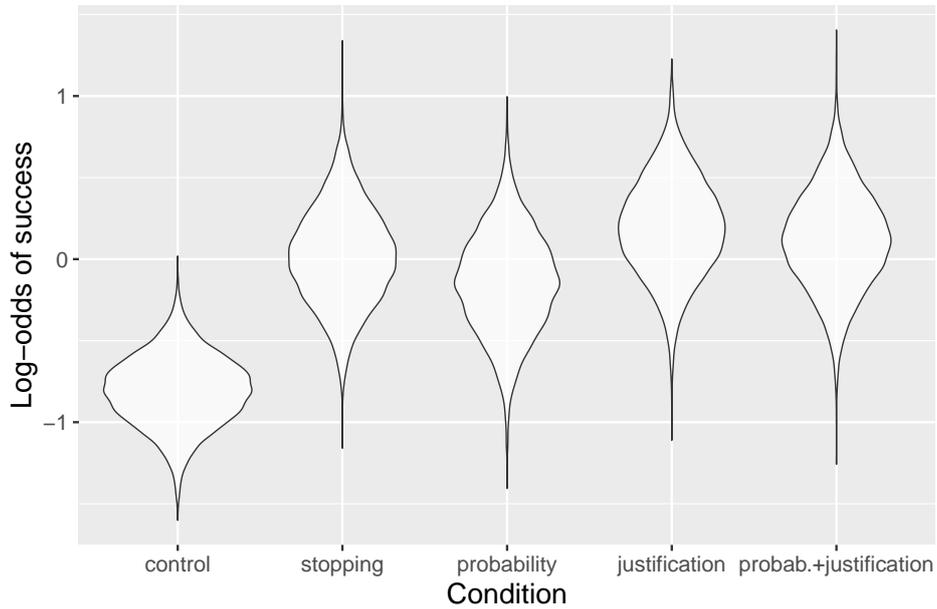


Figure A2. Posterior distributions for the logistic regression of experiment 2.

Appendix A.3. Logistic Regression for Experiment 3

The estimation results of the logistic regression model for experiment 1 are summarized in Figure A3.

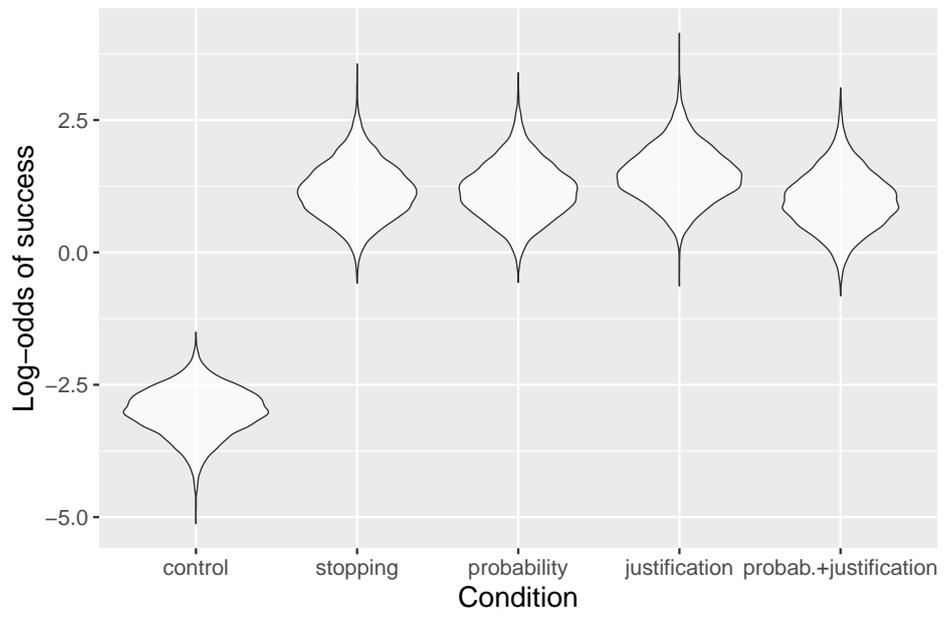


Figure A3. Posterior distributions for the logistic regression of experiment 3.

References

1. Gigerenzer, G. Why Heuristics Work. *Perspect. Psychol. Sci.* **2008**, *3*, 20–29. doi:10.1111/j.1745-6916.2008.00058.x. [[CrossRef](#)] [[PubMed](#)]
2. Gigerenzer, G.; Brighton, H. Homo Heuristicus: Why Biased Minds Make Better Inferences. *Top. Cogn. Sci.* **2009**, *1*, 107–143. doi:10.1111/j.1756-8765.2008.01006.x. [[CrossRef](#)]
3. Simon, H.A. A Behavioral Model of Rational Choice. *Q. J. Econ.* **1955**, *69*, 99–118. doi:10.2307/1884852. [[CrossRef](#)]
4. Kahneman, D. Maps of Bounded Rationality: Psychology for Behavioral Economics. *Am. Econ. Rev.* **2003**, *93*, 1449–1475. [[CrossRef](#)]
5. Lilienfeld, S.O.; Ammirati, R.; Landfield, K. Giving Debiasing Away: Can Psychological Research on Correcting Cognitive Errors Promote Human Welfare? *Perspect. Psychol. Sci.* **2009**, *4*, 390–398. doi:10.1111/j.1745-6924.2009.01144.x. [[CrossRef](#)] [[PubMed](#)]
6. Keren, G. Cognitive Aids and Debiasing Methods: Can Cognitive Pills Cure Cognitive Ills? *Adv. Psychol.* **1990**, *68*, 523–552.
7. Hales, B.M.; Pronovost, P.J. The checklist—A tool for error management and performance improvement. *J. Crit. Care* **2006**, *21*, 231–235. doi:10.1016/j.jcrc.2006.06.002. [[CrossRef](#)]
8. Ely, J.W.; Graber, M.L.; Croskerry, P. Checklists to Reduce Diagnostic Errors. *Acad. Med.* **2011**, *86*, 307. doi:10.1097/ACM.0b013e31820824cd. [[CrossRef](#)]
9. Gawande, A. *The Checklist Manifesto: How to Get Things Right*; Picador: London, UK, 2011.
10. Sibbald, M.; de Bruin, A.B.; van Merriënboer, J.J. Checklists improve experts' diagnostic decisions. *Med. Educ.* **2013**, *47*, 301–308. doi:10.1111/medu.12080. [[CrossRef](#)]
11. Carol-anne, E.M.; Regehr, G.; Mylopoulos, M.; MacRae, H.M. Slowing down when you should: A new model of expert judgment. *Acad. Med. J. Assoc. Am. Med. Coll.* **2007**, *82*, S109–S116. doi:10.1097/ACM.0b013e3181405a76. [[CrossRef](#)]
12. Moulton, C.A.; Regehr, G.; Lingard, L.; Merritt, C.; MacRae, H. Slowing Down to Stay Out of Trouble in the Operating Room: Remaining Attentive in Automaticity. *Acad. Med.* **2010**, *85*, 1571. doi:10.1097/ACM.0b013e3181f073dd. [[CrossRef](#)]
13. Hirt, E.R.; Markman, K.D. Multiple explanation: A consider-an-alternative strategy for debiasing judgments. *J. Personal. Soc. Psychol.* **1995**, *69*, 1069–1086. doi:10.1037/0022-3514.69.6.1069. [[CrossRef](#)]
14. Hirt, E.R.; Kardes, F.R.; Markman, K.D. Activating a mental simulation mind-set through generation of alternatives: Implications for debiasing in related and unrelated domains. *J. Exp. Soc. Psychol.* **2004**, *40*, 374–383. doi:10.1016/j.jesp.2003.07.009. [[CrossRef](#)]
15. Nickerson, R.S. Confirmation bias: A ubiquitous phenomenon in many guises. *Rev. Gen. Psychol.* **1998**, *2*, 175. [[CrossRef](#)]
16. Harman, G.H. The Inference to the Best Explanation. *Philos. Rev.* **1965**, *74*, 88–95. doi:10.2307/2183532. [[CrossRef](#)]
17. Graber, M.L.; Kissam, S.; Payne, V.L.; Meyer, A.N.D.; Sorensen, A.; Lenfestey, N.; Tant, E.; Henriksen, K.; LaBresh, K.; Singh, H. Cognitive interventions to reduce diagnostic error: A narrative review. *BMJ Qual. Saf.* **2012**, *21*, 535–557. doi:10.1136/bmjqs-2011-000149. [[CrossRef](#)]
18. Croskerry, P.; Norman, G. Overconfidence in Clinical Decision Making. *Am. J. Med.* **2008**, *121*, S24–S29. doi:10.1016/j.amjmed.2008.02.001. [[CrossRef](#)] [[PubMed](#)]
19. Dhami, M.K.; Belton, I.K.; Mandel, D.R. The 'Analysis of Competing Hypotheses' in Intelligence Analysis. *Appl. Cogn. Psychol.* **2019**, doi:10.1002/acp.3550. [[CrossRef](#)]
20. Croskerry, P. The Cognitive Imperative: Thinking about How We Think. *Acad. Emerg. Med.* **2000**, *7*, 1223–1231. doi:10.1111/j.1553-2712.2000.tb00467.x. [[CrossRef](#)]
21. Croskerry, P. The Importance of Cognitive Errors in Diagnosis and Strategies to Minimize Them. *Acad. Med.* **2003**, *78*, 775–780. [[CrossRef](#)]
22. Flavell, J.H. Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *Am. Psychol.* **1979**, *34*, 906–911. doi:10.1037/0003-066X.34.10.906. [[CrossRef](#)]
23. Sanna, L.J.; Schwarz, N. Metacognitive Experiences and Human Judgment: The Case of Hindsight Bias and Its Debiasing. *Curr. Dir. Psychol. Sci.* **2006**, *15*, 172–176. doi:10.1111/j.1467-8721.2006.00430.x. [[CrossRef](#)]

24. Schwarz, N.; Sanna, L.J.; Skurnik, I.; Yoon, C. Metacognitive Experiences and the Intricacies of Setting People Straight: Implications for Debiasing and Public Information Campaigns. *Adv. Exp. Soc. Psychol.* **2007**, *39*, 127–161.
25. Frankish, K. Dual-Process and Dual-System Theories of Reasoning. *Philos. Compass* **2010**, *5*, 914–926. doi:10.1111/j.1747-9991.2010.00330.x. [[CrossRef](#)]
26. Evans, J.S.B.T.; Stanovich, K.E. Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspect. Psychol. Sci.* **2013**, *8*, 223–241. doi:10.1177/1745691612460685. [[CrossRef](#)] [[PubMed](#)]
27. Croskerry, P.; Singhal, G.; Mamede, S. Cognitive debiasing 1: Origins of bias and theory of debiasing. *BMJ Qual. Saf.* **2013**, *22*, ii58–ii64. doi:10.1136/bmjqs-2012-001712. [[CrossRef](#)]
28. Helsdingen, A.S.; van den Bosch, K.; van Gog, T.; van Merriënboer, J.J.G. The Effects of Critical Thinking Instruction on Training Complex Decision Making. *Hum. Factors* **2010**, *52*, 537–545. doi:10.1177/0018720810377069. [[CrossRef](#)] [[PubMed](#)]
29. Guess, C.; Donovan, S.; Naslund, D. Improving Dynamic Decision Making through Training and Self-Reflection. *Judgm. Decis. Mak.* **2015**, *10*, 284–295.
30. Fiske, S.T.; Morewedge, C.K.; Yoon, H.; Scopelliti, I.; Symborski, C.W.; Korris, J.H.; Kassam, K.S. Debiasing Decisions: Improved Decision Making With a Single Training Intervention. *Policy Insights Behav. Brain Sci.* **2015**, *2*, 129–140. doi:10.1177/2372732215600886. [[CrossRef](#)]
31. Poos, J.M.; van den Bosch, K.; Janssen, C.P. Battling bias: Effects of training and training context. *Comput. Educ.* **2017**, *111*, 101–113. doi:10.1016/j.compedu.2017.04.004. [[CrossRef](#)]
32. Sherbino, J.; Kulasegaram, K.; Howey, E.; Norman, G. Ineffectiveness of cognitive forcing strategies to reduce biases in diagnostic reasoning: A controlled trial. *Can. J. Emerg. Med.* **2014**, *16*, 34–40. doi:10.2310/8000.2013.130860. [[CrossRef](#)]
33. Dhaliwal, G. Premature closure? Not so fast. *BMJ Qual. Saf.* **2017**, *26*, 87–89. doi:10.1136/bmjqs-2016-005267. [[CrossRef](#)]
34. Croskerry, P. Cognitive forcing strategies in clinical decisionmaking. *Ann. Emerg. Med.* **2003**, *41*, 110–120. doi:10.1067/mem.2003.22. [[CrossRef](#)] [[PubMed](#)]
35. Chew, K.S.; Durning, S.J.; van Merriënboer, J.J. Teaching metacognition in clinical decision-making using a novel mnemonic checklist: An exploratory study. *Singap. Med. J.* **2016**, *57*, 694–700. doi:10.11622/smedj.2016015. [[CrossRef](#)] [[PubMed](#)]
36. Sibbald, M.; Sherbino, J.; Ilgen, J.S.; Zwaan, L.; Blissett, S.; Monteiro, S.; Norman, G. Debiasing versus knowledge retrieval checklists to reduce diagnostic error in ECG interpretation. *Adv. Health Sci. Educ.* **2019**. doi:10.1007/s10459-019-09875-8. [[CrossRef](#)] [[PubMed](#)]
37. Kolodny, N.; Brunero, J. Instrumental Rationality. In *The Stanford Encyclopedia of Philosophy, Winter 2018 ed.*; Zalta, E.N., Ed.; Metaphysics Research Lab, Stanford University: Stanford, CA, USA, 2018.
38. Kahneman, D. Rational Choice and the Framing of Decisions. *J. Bus.* **1986**, *59*, S251–S278.
39. Foley, R. *The Theory of Epistemic Rationality*; Harvard University Press: Cambridge, MA, USA, 1987.
40. Pritchard, D.; Philosophy Documentation Center. Epistemic Luck. *J. Philos. Res.* **2004**, *29*, 191–220. doi:10.5840/jpr_2004_18. [[CrossRef](#)]
41. Lockard, M. Epistemic instrumentalism. *Synthese* **2013**, *190*, 1701–1718. doi:10.1007/s11229-011-9932-6. [[CrossRef](#)]
42. Feldman, R.; Conee, E. Evidentialism. *Philos. Stud.* **1985**, *48*, 15–34. doi:10.1007/BF00372404. [[CrossRef](#)]
43. McCain, K. *Evidentialism and Epistemic Justification*; Routledge: Abingdon-on-Thames, UK, 2014. doi:10.4324/9781315882390.
44. Miller, D. Overcoming the Justificationist Addiction. *Studia Philos. Wratislav.* **2008**, *3*, 9–18.
45. Pappas, G. Internalist vs. Externalist Conceptions of Epistemic Justification. In *The Stanford Encyclopedia of Philosophy*; Metaphysics Research Lab, Stanford University: Stanford, CA, USA, 2014.
46. Alston, W.P. Concepts of Epistemic Justification. *Monist* **1985**, *68*, 57–89. doi:10.5840/monist198568116. [[CrossRef](#)]
47. Foley, R. The Epistemology of Belief and the Epistemology of Degrees of Belief. *Am. Philos. Q.* **1992**, *29*, 111–124.
48. De Finetti, B. Logical foundations and measurement of subjective probability. *Acta Psychol.* **1970**, *34*, 129–145. doi:10.1016/0001-6918(70)90012-0. [[CrossRef](#)]

49. Bovens, L.; Hartmann, S. *Bayesian Epistemol*; Oxford University Press: Oxford, UK, 2004. doi:10.1093/0199269750.001.0001.
50. Talbott, W. Bayesian Epistemology. In *The Stanford Encyclopedia of Philosophy, Winter 2016 ed.*; Zalta, E.N., Ed.; Metaphysics Research Lab, Stanford University: Stanford, CA, USA, 2016.
51. Tversky, A.; Kahneman, D. Belief in the law of small numbers. *Psychol. Bull.* **1971**, *76*, 105–110. doi:10.1037/h0031322. [CrossRef]
52. Bar-Hillel, M.; Wagenaar, W.A. The Perception of Randomness. *Adv. Appl. Math.* **1991**, *12*, 428–454. doi:10.1016/0196-8858(91)90029-I. [CrossRef]
53. Tversky, A.; Kahneman, D. Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychol. Rev.* **1983**, *90*, 293–315. doi:10.1037/0033-295X.90.4.293. [CrossRef]
54. Meehl, P.E.; Rosen, A. Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychol. Bull.* **1955**, *52*, 194–216. [CrossRef]
55. Westbury, C.F. Bayes' Rule for Clinicians: An Introduction. *Front. Psychol.* **2010**, *1*. doi:10.3389/fpsyg.2010.00192. [CrossRef] [PubMed]
56. Vakharia, D.; Lease, M. Beyond AMT: An Analysis of Crowd Work Platforms. *arXiv* **2013**, arXiv:1310.1672.
57. Schmidt, G.B.; Jettinghoff, W.M. Using Amazon Mechanical Turk and other compensated crowdsourcing sites. *Bus. Horiz.* **2016**, *59*, 391–400. doi:10.1016/j.bushor.2016.02.004. [CrossRef]
58. Schmitz, C. LimeSurvey: An Open Source Survey Tool. 2018. Available online: <http://www.limesurvey.org> (accessed on 15 May 2019).
59. Carpenter, B.; Gelman, A.; Hoffman, M.; Lee, D.; Goodrich, B.; Betancourt, M.; Brubaker, M.; Guo, J.; Li, P.; Riddell, A. Stan: A Probabilistic Programming Language. *J. Stat. Softw.* **2017**, *76*, 1–32. doi:10.18637/jss.v076.i01. [CrossRef]
60. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2017.
61. Gelman, A.; Rubin, D.B. Inference from Iterative Simulation Using Multiple Sequences. *Stat. Sci.* **1992**, *7*, 457–472. doi:10.1214/ss/1177011136. [CrossRef]
62. Ghosh, J.; Li, Y.; Mitra, R. On the Use of Cauchy Prior Distributions for Bayesian Logistic Regression. *arXiv* **2015**, arXiv:1507.07170.
63. Gelman, A. Prior Choice Recommendations. Available online: <https://github.com/stan-dev/stan/wiki/Prior-Choice-Recommendations> (accessed on 15 May 2019).
64. Lange, K.; Little, R.J.A.; Taylor, J. Robust Statistical Modeling Using the t-Distribution. *J. Am. Stat. Assoc.* **1989**, *84*, 881–896. [CrossRef]
65. Bürkner, P.C. brms: An R Package for Bayesian Multilevel Models Using Stan. *J. Stat. Softw.* **2017**, *80*, 1–28. doi:10.18637/jss.v080.i01. [CrossRef]
66. Kovic, M. Debiasing in a Minute or Less. 2018. Available online: <https://osf.io/q6wv7/> (accessed on 15 May 2019).
67. Hoekstra, R.; Morey, R.D.; Rouder, J.N.; Wagenmakers, E.J. Robust misinterpretation of confidence intervals. *Psychon. Bull. Rev.* **2014**, *21*, 1157–1164. doi:10.3758/s13423-013-0572-3. [CrossRef] [PubMed]
68. Sanna, L.J.; Schwarz, N.; Stocker, S.L. When debiasing backfires: Accessible content and accessibility experiences in debiasing hindsight. *J. Exp. Psychol. Learn. Mem. Cogn.* **2002**, *28*, 497. [CrossRef] [PubMed]
69. Posch, L.; Bleier, A.; Lechner, C.; Danner, D.; Flöck, F.; Strohmaier, M. Measuring Motivations of Crowdworkers: The Multidimensional Crowdworker Motivation Scale. *arXiv* **2017**, arXiv:1702.01661.
70. Arkes, H.R. Costs and benefits of judgment errors: Implications for debiasing. *Psychol. Bull.* **1991**, *110*, 486–498. doi:10.1037/0033-2909.110.3.486. [CrossRef]
71. Willems, S.J.W.; Albers, C.J.; Smeets, I. Variability in the interpretation of Dutch probability phrases—A risk for miscommunication. *arXiv* **2019**, arXiv:1901.09686.

