Journal of Behavioral Decision Making

WILEY

RESEARCH ARTICLE

Blaming the Strawless Brickmaker: Constraint Neglect in Judging Decision Quality

Xilin Li[1] | Christopher K. Hsee[2]

[1]Marketing Department, China Europe International Business School (CEIBS), Shanghai, China | [2]Marketing Department, Cheung Kong Graduate School of Business, Beijing, China

Correspondence: Xilin Li (xilinli@ceibs.edu)

Received: 22 July 2024 | **Revised:** 4 November 2024 | **Accepted:** 10 December 2024

Keywords: decision quality | hiring decision | judgment bias | performance evaluation

ABSTRACT

People often judge the quality of selection decisions made by others: The CEO of a firm may judge the quality of hiring decisions made by the firm's HR personnel; the readers of a journal may judge the quality of manuscript-acceptance decisions by the journal's editor. To accurately judge others' selection decision quality, evaluators should consider not only the outcome of the selection decisions but also the constraints of the decision-maker. For example, to judge the quality of the hiring decisions made by the HR personnel, the CEO should consider not only how many high-quality (vs. low-quality) candidates the HR personnel hired, but also how many high-quality (vs. low-quality) candidates applied, and how many candidates the HR personnel needed to hire. We theorize that evaluators tend to overlook these constraints, and, consequently, judge decision-makers who faced greater constraints as having made worse decisions than decisions-makers who faced lesser constraints, even if the former's decisions were actually as good as or better than the latter's. We refer to this phenomenon as Blaming the Strawless Brickmaker (from the saying "making bricks without straw"). Eight studies, employing mixed methods, demonstrate the Blaming-the-Strawless-Brickmaker effect, examine its underlying mechanism, and suggest ways to improve the quality of judged selection decision quality.

Undoubtedly, making high-quality decisions is crucial for individuals and organizations. Extensive existing research studied factors that influence decision quality (e.g., Dane, Rockmann, and Pratt 2012; Derfler-Rozin and Pitesa 2020; Dooley and Fryxell 1999; Lerner et al. 2015; Li and Hsee 2019a, 2019b; Lurie and Swaminathan 2009; O'Reilly 1982; Seo and Barrett 2007) and ways to improve decision quality (e.g., Arkes 1991; Fisher 2017; Sellier, Scopelliti, and Morewedge 2019; Yoon, Scopelliti, and Morewedge 2021).

Instead of studying decision quality per se, the current paper studies the *judgment* of decision quality, especially the judgment of the quality of the decisions made by others. Understanding how people judge others' decision quality is valuable, because accurate judgments lead to fair treatment of the decision-makers, and inaccurate judgments lead to mistreatments in organizations.

People often judge others' decision quality: The CEO of a company may judge the quality of the hiring decisions made by the HR personnel, the deans of a college may judge the quality of the admissions decisions made by the admissions office, and the readers of a journal may judge the quality of the manuscript-acceptance decisions made by the journal's editor. These judgments may affect the promotions or demotions of the decision-maker and the future direction of the institutions.

To accurately judge someone else's decision quality, the evaluator should consider not only the output quality of the decisions, but also the constraints of the decision-maker. For example, to

judge the quality of the manuscript-acceptance decisions made by the editor of a journal, readers should consider not only how many high-quality versus low-quality manuscripts the editor has accepted but also how many high-quality versus low-quality manuscripts were submitted to the journal in the first place and how many manuscripts the editor had to accept in order to fill the space of the journal.

Note that the decisions in our current research pertain specifically to selection decisions. These decisions involve selecting target objects from a larger pool. To elaborate, constraints in our study refer to the discrepancy between the number of high-quality candidates available in the pool and the number of candidates actually needed. Additionally, decision quality denotes the efficacy in selecting high-quality candidates over low-quality candidates from this pool.

We theorize that evaluators tend to focus on the output quality of the decisions, and overlook the constraints of the decision-maker. As a result, a decision-maker who encountered greater constraints will be judged as having made worse decisions than a decision-maker who encountered lesser constraints, even though the former has actually made as good decisions as, or even better decisions, than the latter.

A main reason why real-life evaluators may overlook the constraints of the decision-maker is that such information is not available to them. For example, most readers of a journal do not know how many high-quality versus low-quality manuscripts were submitted, or how many manuscripts the editor had to accept; they only see how many high-quality vs. low-quality manuscripts are eventually published.

In this research, we examine situations in which information about the constraints of the decision-maker is readily available to the evaluator. We test whether evaluators in such situations still overlook the information and render unfair judgments. If so, then it strongly suggests that evaluators in real life, who lack such information, are even more likely to render unfair judgments. In the ensuring sections, we elaborate on our theory, propose specific hypotheses, and report empirical studies.

1 | Theory

Consider the following stylized case: The CEO of a company posted recruitment ads to hire new employees. Among the people who responded to the ads and applied, some were high-quality and some were low-quality. (For ease of analysis, we assume that the quality of a candidate is either high or low. In Section 11, we discuss situations in which the quality of candidates is a continuous variable.) The CEO asked an assistant to make hiring decisions. The assistant had no control over the number or the quality of the candidates; the only decisions they could make were to decide which candidates to hire. Based on the needs of the company, the CEO requires the assistant to hire a specific number of candidates. The assistant followed the CEO's instructions, and hired the required number of candidates.

1.1 | Objective Decision Quality Versus Judged Decision Quality

In this section, we first discuss what unbiased evaluators *should* do, and then theorize what ordinary evaluators *actually* do.

1.1.1 | What Unbiased Evaluators Should Do

To answer this question, we need to define two terms—Best Possible Output Quality and Realized Output Quality. Best Possible Output Quality is the best outcome that the decision-maker can possibly produce given the constraints they face. Formally,

$$\textit{Best Possible Output Quality} = \frac{min(\#\textit{InHQ}, \#\textit{Needed})}{\#\textit{Needed}} \quad (1)$$

where #InHQ is the number of high-quality candidates in the candidate pool and #Needed is the number of candidates the decision-maker is obligated to accept given the needs in the situation. The nominator in the equation—min(#Needed, #InHQ)—means the smaller of #Needed and #InHQ; it reflects two intertwined constraints the decision-maker faces: (a) the decision-maker cannot accept more high-quality candidates than the number of high-quality candidates available in the candidate pool (i.e., #InHQ) and (b) the decision-maker must accept the needed number of candidates (i.e., #Needed) even if the number of high-quality candidates in the candidate pool is smaller than the needed number of candidates.[1]

Unlike Best Possible Output Quality, Realized Output Quality is the actual outcome realized by the decision-maker, which we define as the proportion of high-quality candidates among the candidates actually accepted by the decision-maker, namely,

$$\textit{Realized Output Quality} = \frac{\#\textit{OutHQ}}{\#\textit{Needed}} \quad (2)$$

In the equation, the nominator—#OutHQ—is the number of high-quality candidates accepted by the decision-maker, and the denominator is the total number of candidates accepted by the decision-maker, which is the same as #Needed, because the decision-maker must accept the needed number of candidates.

With the above terms, we now answer the question posed earlier: What unbiased evaluators *should* consider when judging the decision quality of the decision-maker. We argue that an unbiased evaluator should consider how close the decision-maker's Realized Output Quality is to the Best Possible Output Quality. The closer (farther) it is, the better (worse) their decision quality is. For ease of exposition, we will use the term Objective Decision Quality to denote the proximity between the two variables, namely, the signed difference between Realized Output Quality and Best Possible Output Quality. Formally,

$$\textit{Objective Decision Quality} = (\textit{Realized Output Quality} - \textit{Best Output Quality}) + 1 \quad (3)$$

(The "+1" at the end is merely to make the range of Objective Decision Quality consistent with the range of Realized Output Quality and the range of Best Output Quality—all between 0 and 1.)

The rationale behind Equation (3) is that in order to judge one's decision quality, the evaluator should consider whether the decision-maker has done the best they could possibly do; namely, the evaluator should use Best Possible Output Quality as the benchmark to judge their Realized Output Quality.[2] The equation aligns with classic selection decision research showing that both the proportion of satisfactory employees and selection ratio are critical factors (Taylor and Russell 1939). It also reflects Brogden and Taylor's (1950) argument that performance evaluation must consider criteria contamination—factors beyond workers' control that significantly affect their performance.

We are not saying that Objective Decision Quality is the only factor that an unbiased evaluator should consider, but it is the most important factor. In General Discussion, we discuss other potential factors we think an unbiased evaluator should consider.

1.1.2 | What Evaluators Actually Do

In the above section, we discussed what unbiased evaluators *should* do when assessing the decision quality of a decision-maker. Now, we discuss what (ordinary) evaluators actually do. This is our main research question.

We propose that evaluators underweight Best Possible Output Quality. Formally, we propose:

$$\text{Judged Decision Quality} = (\text{Realized Output Quality} - w \times \text{Best Possible Output Quality}) + w \quad (4)$$

The key difference of Equations (4) from (3) is the inclusion of the weight *w*. If *w* were 1, Equation (4) would be identical to Equation (3).

The crux of our proposition is that w is less than 1; this is what we mean by saying that evaluators "underweight" Best Possible Output Quality. In other words, they underuse Best Possible Output Quality as the benchmark for judging Realized Output Quality. (In the extreme case, $w = 0$, and Equation (4) will simply be Judged Decision Quality = Realized Output Quality.)

Recall that Best Possible Output Quality is composed of the two constraint variables--#InHQ and #Needed (Equation (1)). Therefore, to say that evaluators underweight Best Possible Output Quality means that they underweight these two constraint variables. We refer to this tendency as *constraint neglect*. More specifically, we refer to the tendency to underweight #InHQ as *input neglect*, and the tendency to underweight #Needed as *need neglect*. Input neglect and need neglect are two manifestations of constraint neglect.

Constraint neglect does not mean that people completely disregard constraints and give zero weight to the information to the factor; rather, we suggest that they do not pay sufficient attention to these constraints when evaluating selection decision quality. This understanding aligns with existing literature, which indicates that people often overlook factors that should normatively be considered. Classic examples include base rate neglect (Bar-Hillel 1980; Kahneman 2011), duration neglect (Fredrickson and Kahneman 1993; Liersch and McKenzie 2009), scope neglect (Frederick and Fischhoff 1998; Hsee and Rottenstreich 2004), opportunity cost neglect (Frederick et al. 2009), and marginal utility neglect (Li and Hsee 2021b). For example, duration neglect does not imply that people completely disregard the duration when evaluating experiences; rather, it means that they are not adequately sensitive to the duration. Similarly, opportunity cost neglect does not mean that people entirely overlook opportunity costs; it suggests that they do not assign sufficient weight to them.

1.2 | Related Literature

Our constraint-neglect theory is inspired by multiple lines of existing research. One is classic literature on the correspondence bias (Gilbert and Malone 1995; Moore et al. 2010; Ross 1977; Scopelliti et al. 2018). When explaining an action conducted by someone else, people tend to attribute the action to the actor's stable internal factors (e.g., personality) and overlook uncontrollable external situational factors. For example, if a student turned in an assignment late, the teacher is likely to consider the student lazy and overlook the possibility that the student may have been infected with Covid and sick. Our current research corroborates the existing literature by showing that the evaluator overlooks an important type of situational factor—the constraints of the decision-maker. However, our research is concerned with the judgment of decision quality rather than the attribution of an action.

Another existing finding that has inspired our work is the outcome bias—the phenomenon that people rely on the ex post outcome of a decision to judge the ex ante correctness of the decision, and overlook the effect of random factors on the outcome (Baron and Hershey 1988; Savani and King 2015; Lefgren, Platt, and Price 2015). For example, in a study by Baron and Hershey (1988), participants were told that a decision-maker who had faced a choice between a sure gain and a gamble with a probabilistic gain decided to choose the gamble and were asked whether it was a good choice. The participants judged the choice as good if the outcome was good, or bad if the outcome was bad, even though the participants were told that the outcome of the gamble was determined by the spin of a wheel. Apparently, the participants overlooked the influence of random factors on the outcome. Our research is consistent with this line of research by revealing that when judging the quality of a decision, people over-rely on the outcome of the decision and overlook factors that the decision-maker had no control over. However, our present research is concerned with skilled-based decisions (rather than stochastic decisions) or the oversight of objective constraints (rather than random factors).

A third line of existing literature that has inspired our work is the base rate neglect. When judging the likelihood of an event, people tend to focus on the event-specific probability and overlook the base rate of the event (Bar-Hillel 1980; Kahneman and

Tversky 1973; Welsh and Navarro 2012). For example, when judging the likelihood that a certain MBA student in a business school had a college degree in classic literature, people over-rely on how much the student resembles a typical classic-literature major and overlook the proportion of classic-major students in the business school. Our research is consistent with this line of research by revealing that people over-rely on the outcome of the decision, and overlook factors that the decision-maker had no control over. Our research differs from base-rate neglect in three aspects: judgment target, weighted factors, and mathematical requirements. First, while base-rate neglect concerns probability judgments with updated information, our research examines decision-quality judgments. Second, base-rate neglect involves underweighting prior probabilities and overweighting case-specific information, whereas our research shows underweighting of selection constraints and overweighting of realized outputs. Third, base-rate neglect requires strong mathematical skills to calculate Bayesian posterior probabilities, while evaluating selection decision quality is relatively straightforward and requires minimal mathematical ability.

At a more general level, our work is consistent with a rich body of literature showing that people tend to base their judgment and choice on salient "foreground information," and overlook important "background information" (Kahneman 2011). For example, when judging the quality of a person (e.g., how good a certain college student is), people tend to focus on the rank of the person in their category which is salient "foreground information" (e.g., how good the student is in her college) and overlook the quality of the category per se which is non-salient (background information) (e.g., how good the college is) (Leclerc, Hsee, and Nunes 2005; Read, Loewenstein, and Rabin 1999). Like these findings, our research shows that when judging the quality of someone's decisions, people focus on the "foreground information"—the output quality of the decisions and overlook the "background information"—the constraints of the decision-maker. In other words, our contribution does not entail the creation of a new theory; rather, it involves applying an existing theoretical framework to a novel and significant context.

Although the present research is inspired by multiple lines of existing research, it makes unique contributions. No existing research has studied the constraint-neglect issue we study in this research. The constraint-neglect issue studied in this research is as distinct from any of the above-reviewed projects as the above-reviewed projects are distinct from each other. For example, at a general level, the correspondence bias and base-rate neglect are related to each other, because both biases reflect an oversight of important background information. However, those two biases concern different background information (situational factor vs. base rate), and different dependent variables (attribution vs. probability judgment), so neither of them subsumes the other. Likewise, at a general level, the constraint-neglect bias studied in the current research also reflects an oversight of important background information, but our research concerns different background information (the constraints of the decision-maker) and a different dependent (the judgment of decision quality) than the correspondence bias and base rate neglect.

From a practical point of view, the constraint-neglect problem investigated in this research is important, because it is widespread in real-life decision situations: Many decision-makers, be they HR personnel, admissions officers, or journal editors, face such constraints. This research sheds light on how evaluators judge the decision-makers in these commonly-encountered situations and can potentially help them make more accurate and fairer judgments.

From a theoretical point of view, our constraint-neglect theory does not just make an general statement that evaluators overlook constraint-related information; rather, it makes specific and testable hypotheses, which we elaborate on in the next section.

1.3 | Specific Hypotheses

1.3.1 | Primary Hypothesis

Before introducing the hypothesis, we first introduce the empirical paradigm of our research, as we will present the hypothesis in the context of this paradigm. In the paradigm, research participants are the evaluators, and their task is to judge the decision qualities of two decision-makers: a more-constrained decision-maker and a less constrained decision-maker. We operationally define the more-constrained and the less-constrained decision-makers as such that the Best Possible Output Quality achievable by the more-constrained decision-maker is lower than the Best Possible Output Quality achievable by the less-constrained decision-maker. In our studies, we manipulate Best Possible Output Quality by varying one or both of the two constraint variables—#InHQ and #Needed. Unless otherwise specified, the evaluators (research participants) know the constraints of the two decision-makers, and also know their output quality. Therefore, the evaluators have all the relevant information to infer their Objective Decision Quality.

In the context of this empirical paradigm, our constraint-neglect theory predicts the following:

H1. : *Evaluators will judge the more-constrained decision-maker as having made worse decisions than the less-constrained decision-maker, even though the former's Objective Decision Quality is as high as or higher than the latter's, and the evaluator has all the relevant information to infer that.*

This is our primary hypothesis, and we dub the hypothesized effect as *Blaming the Strawless Brickmaker*, from the idiom "making bricks without straw" (https://en.wikipedia.org/wiki/Bricks_without_straw). The more-constrained decision-maker is a strawless brickmaker. Just as the strawless brickmaker has to make bricks without sufficient straw, the more-constrained decision-maker has to accept a given number of candidates without sufficient high-quality candidates.

1.3.2 | Secondary Hypotheses

Our primary hypothesis (H1) is about information usage, positing that evaluators "underuse" (underweight) the constraint-related information in their judgment of decision quality. Our second hypothesis (H2) is about information search. We

propose that evaluators not only underweight constraint-related information when they have the information, they also underweight constraint-related information if they do not have it but can expend resources to acquire it. Specifically, we propose that if the evaluator has neither the constraint-related information nor the output-quality information but can expend resources to acquire the information, they are less likely to seek the constraint-related information than to seek the output-quality information. This is our second hypothesis (H2):

H2. : *During the information search phase, evaluators are less likely to seek information about the constraints of the decision-makers than to seek information about their output quality.*

In summary, both H1 and H2 show constraint neglect. Whereas, H1 shows constraint neglect in the use of available information, H2 shows constraint neglect in the search for not-yet-available information.

Besides H1 and H2, another way to test our constraint-neglect theory is to examine whether prompting the evaluator to consider the constraints will moderate the Blaming-the-Strawless-Brickmaker effect. If the Blaming-the-Strawless-Brickmaker effect is indeed due to the neglect of the constraints as our theory posits, then prompting the evaluator to consider the constraints will moderate the effect. This is our last hypothesis:

H3. : *The Blaming-the-Strawless-Brickmaker effect will attenuate if the evaluator is prompted to consider the constraints of the decision-makers.*

If H3 proves true, it not only supports the constraint-neglect theory but also addresses an alternative explanation—that people lack mathematical skills. If this effect were simply due to poor mathematical skills, the prompting manipulation would not have a moderating effect, as it provides no new mathematical information nor improves mathematical ability. Our theory suggests that the Blaming-the-Strawless-Brickmaker effect stems from insufficient deliberation rather than mathematical inability: People can perform the calculations but fail to consider constraints. Therefore, the prompting manipulation can effectively reduce this effect. Moreover, if H3 is supported, the prompting manipulation could serve as a strategy to improve judgment accuracy in evaluating decision quality.

2 | Study Overview

We next report eight studies that tested our theory. Because H1 (Blaming the Strawless Brickmaker) is our primary hypothesis, is most directly derived from our constraint-neglect theory and has never been tested before, we find it important to test this hypothesis across multiple studies to ensure that the hypothesized effect is replicable and robust. Therefore, we devoted six of our eight studies (Study 1A through Study 1F) to testing this hypothesis. These studies manipulated constraints of the decision-makers in different ways (by varying #InHQ, by varying #Needed, or by varying both), involved participants with different backgrounds and from different cultures (online workers in North America and Europe and people with managerial experiences in Asia), and adopted different methods (scenario-based judgment and incentive-compatible choice). The results lent strong support to H1 and the constraint-neglect theory and ruled out alternative explanations. In addition to the six studies testing our primary hypothesis, we also conducted two studies to test our secondary hypotheses: Study 2 tested H2 regarding information search, and Study 3 tested H3 regarding the moderating effect of prompting participants to consider constraints.

We have conducted an a-priori power analysis for the statistical tests involved in our analysis and set a target sample size of 150 participants per condition, which is larger than the minimum sample sizes calculated based on these power analyses ($\alpha = 0.05$, $power = 0.80$).

3 | Study 1A

Study 1A tested H1 regarding the Blaming-the-Strawless-Brickmaker effect. Participants read a case about two decision-makers—Abby (the more-constrained decision-maker) and Barb (the less-constrained decision-maker)—and judged their respective decision qualities. We varied the constraints of the two decision-makers by varying #InHQ, while holding #Needed constant.

3.1 | Method

The study was pre-registered at https://aspredicted.org/VMJ_NKT. Participants were 150 online workers recruited from Prolific (71 females, $M_{age} = 27.89$). They read the following case:

> The CEO of a company recently posted recruitment ads to recruit employees for two newly-created departments—A and B. A total of 40 people responded to the recruitment ads and applied. Half of them applied to department A and the other half applied to department B. Among those who applied to department A, 4 were high-quality and 16 were low-quality; among those who applied to department B, 8 were high-quality and 12 were low-quality.

> The CEO asked two assistants—Abby and Barb—to make hiring decisions (i.e., identifying and hiring high-quality applicants from the applicant pool). The CEO assigned Abby to make hiring decisions for department A and Barb to make hiring decisions for department B. The assignment was random and did not reflect the relative ability of the two assistants.

> The CEO required Abby to hire 10 applicants for department A and Barb to hire 10 applicants for department B. Abby and Barb followed the CEO's instructions and hired the required numbers of applicants for their respective departments.

> Among the applicants Abby hired, 4 were high-quality and 6 were low-quality; among the applicants Barb hired, 8 were high-quality and 2 were low-quality.

Participants were then asked to judge the quality of Abby's decisions and the quality of Barb's decisions. For each judgment, participants were given six options—extremely low, quite low, somewhat low, somewhat high, quite high, and extremely high. In our analyses, we converted participants' responses to a 1–6 scale, where *extremely low* = 1 and *extremely high* = 6.

3.2 | Results and Discussion

Table 1 summarizes the parameters and the results of the study. As the table shows, among the candidates available for Abby to hire, only four were high-quality, while among the candidates available for Barb to hire, eight were high-quality, and both of them were required to hire 10 candidates. Therefore, Abby was more constrained than Barb in the sense that Abby had less "straw" than Barb but was required to make as many "bricks" as Barb.

Both Abby and Barb did their best—hired all the high-quality candidates available for them to hire, so that their Objective Decision Quality was equally high. However, the participants showed our hypothesized Blaming-the-Strawless-Brickmaker effect; judging the quality of Abby's decisions to be significantly lower than that of Barb's decisions, $t(149) = 5.13$, $p < 0.001$, Cohen's $d = 0.45$. Apparently, participants based their judgments on the Realized Output Quality of the decision-makers and overlooked the difference in the constraints they had faced, specially, the difference in #InHQ in their respective candidate pool.

4 | Study 1B

Study 1B was a replication of Study 1A. Instead of varying #InHQ while holding #Needed constant, Study 1B varied #Needed while holding #InHQ constant.

4.1 | Method

The study was pre-registered at https://aspredicted.org/5Y1_NQ8. Participants were 151 online workers recruited from Prolific (76 females, $M_{age} = 27.68$). The procedure was the same as in Study 1A, except that the parameters were different, as show in Table 2.

4.2 | Results and Discussion

As the top section of Table 2 shows, there were 6 high-quality candidates available for each decision-maker to hire, but Abby

TABLE 1 | Parameters and results of Study 1A.

	More-constrained decision-maker (Abby)	Less-constrained decision-maker (Barb)
#InHQ	4	8
#Needed	10	10
Best Possible Output Quality	4/10	8/10
Realized Output Quality	4/10	8/10
Objective Decision Quality	1	1
Judged Decision Quality (mean and SD)	4.34 (1.37)	4.88 (1.00)

TABLE 2 | Parameters and results of Study 1B.

	More-constrained decision-maker (Abby)	Less-constrained decision-maker (Barb)
#InHQ	6	6
#Needed	16	8
Best Possible Output Quality	6/16	6/8
Realized Output Quality	6/16	6/8
Objective Decision Quality	1	1
Judged Decision Quality (mean and SD)	4.34 (1.34)	4.96 (0.93)

was required to hire 16 candidates, while Barb was required to hire only 8 candidates. Therefore, Abby was more constrained than Barb, in the sense that with the same amount of "straw," Abby had to "make more bricks" than Barb had to. Both Abby and Barb did their best—hired all the high-quality candidates available for them to hire, so that their Objective Decision Quality was equally high.

However, as the last row of Table 2 reports, the participants again showed the Blaming-the-Strawless-Brickmaker effect; judging the quality of Abby's decisions to be significantly lower than that of Barb's decisions, $t(150) = 6.54$, $p < 0.001$, Cohen's $d = 0.54$. Apparently, participants based their judgments on the Realized Output Quality of the decision-makers, and overlooked their difference in #Needed.

Both Study 1A and Study 1B supported our constraint-neglect theory, but they supported different aspects of the theory: Study 1A supported the input-neglect of the theory (overlooking the number of high-quality candidates available for each decision-maker to hire), and Study 1B supported the need-neglect aspect of the theory (overlooking the number of candidates each decision-maker needed to hire).

5 | Study 1C

Study 1C addressed two potential alternative explanations for the Blaming-the-Strawless-Brickmaker effect observed in Study 1A and Study 1B. One is that the effect occurred, not because participants overlooked constraints, but because they found Barb's outcome more difficult to be achieved, namely, less likely to be achieved merely by luck (chance). Take Study 1A, for example. Statistically, it is less likely for Barb to have achieved what she achieved (i.e., picked eight high-quality candidates from a pool with eight high-quality candidates) merely by chance than for Abby to have achieved what she achieved (i.e., picked four high-quality candidates from a pool with four high-quality candidates) merely by chance. Therefore, one may argue that Barb must have made better decisions. To address this alternative explanation, we designed Study 1C so that this alternative explanation would make the opposite prediction to our constraint-neglect theory (see below for details).

A second potential alternative explanation for the Blaming-the-Strawless-Brickmaker effect found in Study 1A and Study 1B is that the effect occurred, not because participants did not know that Abby had made better decisions, but because they misunderstood what we wanted them to judge and (mistakenly) thought that we merely wanted them to judge Realized Output Quality (i.e., the proportion of high-quality candidates among the accepted candidates). To address this concern, Study 1C not only asked the participants to judge the decision quality of each decision-maker, but also asked them which decision-maker they would entrust to make decisions in the future. If participants were more likely to entrust the less-constrained decision-maker than to entrust the more-constrained decision-maker to make future decisions (holding their Objective Decision Quality constant), then this is behavioral evidence that the participants must have deemed the decision quality of the less-constrained decision-maker as better.

5.1 | Method

Participants were 152 online workers recruited from Prolific (87 females, $M_{age} = 30.50$). They read the following case:

> You are the owner of a company. You recently posted recruitment ads to hire employees for two newly-created departments—department A and department B.
>
> You were too busy to make hiring decisions yourself, and so you asked your assistants Alex and Bob to make hiring decisions for you. You assigned Alex to make hiring decisions for department A and Bob to make hiring decisions for department B. The assignment was random and did not reflect your belief about their relative ability. Because they were assigned to make hiring decisions after all applications had come in, they had no control over the number or quality of the applicants; the only decisions they could make were to decide which applicants to hire.
>
> You are less busy now, and you want to find out how good their hiring decisions were. You have gathered the following information:
>
> After you had posted the recruitment ads, a total of 100 people responded and applied. Half of them applied to department A and the other half applied to department B. You check the quality of the applicants, and find that among those who applied to department A, 5 were high-quality and 45 were low-quality, and among those who applied to department B, 45 were high-quality and 5 were low-quality.
>
> Based on the needs of the two new departments, you asked each of Alex and Bob to hire 20 applicants from their respective applicant pools for their respective departments. They followed your instructions and hired the required number of applicants.
>
> You check the quality of the applicants Alex and Bob hired, and find that among those Alex hired, 5 were high-quality and 15 were low-quality, and among those Bob hired, 20 were high-quality and 0 were low-quality.

Participants were then asked two questions. The first concerned decision quality—"Who made better hiring decisions?" The second concerned future decisions—"You now want to ask one of them to make hiring decisions for another newly-created department. Whom would you ask?" Participants answered each question on a 7-point scale, anchored by "Definitely Alex" and "Definitely Bob."

5.2 | Results and Discussion

As the top part of Table 3 shows, there were only five high-quality candidates in Alex's pool and there were 45 high-quality candidates in Bob's pool, and both were required to hire 20 candidates. So, Alex faced greater constraints than Bob. Given their respective constraints, both Alex and Bob did their best, so their Objective Decision Quality was the same.

Statistically, it is less likely for Alex to have achieved what he achieved (i.e., picked five high-quality candidates from a pool with five high-quality candidates) merely by chance than for Bob to have achieved what he achieved (i.e., picked 20 high-quality candidates from a pool with 45 high-quality candidates) merely by chance.[3] In this regard, what Alex achieved was more difficult than what Bob achieved. If participants based their judgment of decision quality on decision difficulty, they should have judged Alex as having made better decisions.

However, as the bottom part of Table 3 displays, our participants (on average) judged Bob as having made better decisions than Alex, $t(151) = 4.54$, $p < 0.001$, Cohen's $d = 0.93$, when compared with the neutral point of the decision-quality scale. Specifically, only a minority of the participants (31.6%) judged the two decision-makers as having made equally good decisions (i.e., gave the neutral rating). Significantly more participants judged Bob as having made better decisions than participants who judged Alex as having made better decisions, 45.5% versus 23.0%, $\chi^2 = 11.12$, $p = 0.004$. These results were consistent with our theory, and inconsistent with the decision-difficulty account.

Not only did participants judge Bob as having made better decisions, they were also more likely to entrust Bob to make future decisions, $t(151) = 4.26$, $p < 0.001$, Cohen's $d = 0.71$, when compared with the neutral point of the future-decision scale. Specifically, only a minority of the participants (20.2%) were indifferent between the two decision-makers (i.e., gave the neutral rating). Significantly more participants wanted Bob to make future decisions for them than participants who wanted Alex to make future decisions for them, 52.0% versus 27.0%, $\chi^2 = 12.06$, $p = 0.002$, xx.

These results replicated the Blaming-the-Strawless-Brickmaker effect in both judged decision quality and delegation of future decisions and cast double on decision difficulty and misunderstanding as alternative explanations.

6 | Study 1D

Study 1D extended the studies reported above in two main aspects. First, while the studies reported so far tested H1 *within-subjects*, Study 1D tested it *between-subjects*. We asked half the participants to judge the decision quality of a more-constrained decision-maker and the other to judge the decision quality of a less-constrained decision-maker. We ran this between-subjects replication for both a practical reason and a theoretical reason. Practically, evaluators in real life often judge the decision quality of a single decision-maker (rather than multiple decision-makers), and the between-subjects paradigm reflects this reality. Theoretically, previous research shows that preferences elicited in joint evaluation (i.e., within-subject elicitation) may disappear or reverse in separate evaluation (i.e., between-subjects elicitation) (Bazerman, Loewenstein, and White 1992; Hsee 1996), due to differential evaluability or differential justifiability (Hsee and Zhang 2010; Li and Hsee 2019a, 2019b; Li and Hsee 2021a, 2021b). However, the constraints we study in this research do not involve differential evaluability or differential justifiability. Therefore, theoretically, we expected the Blaming-the-Strawless-Brickmaker effect to replicate in separate evaluation. Study 1D empirically tested this expectation.

Second, in the previous studies, both the more-constrained and the less-constrained decision-makers did their best given their respective constraints, but in this study, neither decision-maker did their best given their respective constraints. This feature also reflects many real-life situations.

6.1 | Method

Participants were 301 online workers recruited from Prolific (123 females, $M_{age} = 30.40$). They were randomly assigned to two conditions—more-constrained and less-constrained.

TABLE 3 | Parameters and results of Study 1C.

	More-constrained decision-maker (Alex)	Less-constrained decision-maker (Bob)
#InHQ	5	45
#Needed	20	20
Best Possible Output Quality	5/20	20/20
Realized Output Quality	5/20	20/20
Objective Decision Quality	1	1
Judged Decision Quality (mean and SD)	−0.72 (1.96)	+0.72 (1.96)
Delegation to make future decisions	−0.71 (2.06)	+0.71 (2.06)

Note: The results on the last two rows are differences from the neutral point of the respective scales (i.e., 4). Because both scales are relative, the differences are symmetrical for the two decision-makers.

Specifically, participants in the more-constrained condition read the following:

> You are the owner of a company. You recently posted recruitment ads to hire employees for a new department. Many people responded and applied.
>
> You were too busy to make hiring decisions yourself, and so you asked your assistant Dan to make hiring decisions. (Dan had no control over the number or quality of the applicants; the only decisions he could make were to decide which applicants to hire.) Based on the size of the new department, you asked Dan to hire 30 people from the applicant pool, and he did.
>
> You are less busy now, and you want to find out how good Dan's hiring decisions were. You have gathered the following information:
>
> After you had posted the recruitment ads, a total of 40 people responded and applied. Among them, 10 were high-quality and 30 were low-quality.
>
> You asked Dan to hire 30 applicants; he followed your instructions and hired 30 applicants. Among the applicants Dan hired, 9 were high-quality and 21 were low-quality.

The less-constrained condition was identical to the more-constrained condition, except that the number of high-quality candidates in the candidate pool was changed from 10 to 30, and the number of candidates the decision-maker needed to hire was changed from 30 to 10. Accordingly, the last part of the stimuli in the less-constrained condition reads as follows:

> After you had posted the recruitment ads, a total of 40 people responded and applied. Among them, 30 were high-quality and 10 were low-quality.
>
> You asked Dan to hire 10 applicants; he followed your instructions and hired 10 applicants. Among the applicants Dan hired, 9 were high-quality and 1 was low-quality.

Finally, all participants were asked "How good were Dan's hiring decisions?" They answered on a 9-point scale, anchored by "not very good" (1) and "extremely good" (9).

6.2 | Results and Discussion

As the top part of Table 4 summarizes, the decision-maker in the more-constrained condition had only 10 high-quality candidates available for hire, but had to hire 30 candidates, while the decision-maker in the less-constrained condition had 30 high-quality candidates available for hire, but had to hire only 10 candidates. In both conditions, the decision-maker did not do their best: They hired only 9 out of 10 high-quality candidates that he could have hired. Notably, the Objective Decision Quality of the decision-maker in the more-constrained condition was actually higher than that of the decision-maker in the less-constrained condition.

TABLE 4 | Parameters and results of Study 1D.

	More-constrained condition	Less-constrained condition
#InHQ	10	30
#Needed	30	10
Best Possible Output Quality	10/30	10/10
Realized Output Quality	9/30	9/10
Objective Decision Quality	0.97	0.90
Judged Decision Quality (mean and SD)	6.35 (2.12)	7.22 (1.54)

Nevertheless, our participants judged the decision-maker in the more-constrained condition significantly less favorably than the decision-maker in the less-constrained condition, $t(299) = 4.07$, $p < 0.001$, Cohen's $d = 0.47$. This finding replicated the Blaming-the-Strawless-Brickmaker effect in a between-subjects paradigm, confirming our expectation that the effect is robust to both joint evaluation (within-subjects judgment) and separate evaluation (between-subject judgments).

7 | Study 1E

Study 1E was another replication and extension. While the participants in the studies reported so far were regular online workers in the West, the participants in Study 1E were managers in Asia (China). We wanted to test whether the Blaming-the-Strawless-Brickmaker effect is robust to individuals with different professional and different cultural backgrounds.

In addition, Study 1E addressed a potential concern with the studies reported so far—that the participants did not pay attention to or did not understand the information about the constraints. To minimize this concern, Study 1E required every participant to pass a series of comprehension questions regarding the information.

Like Study 1C, Study 1E not only asked participants to judge decision quality; it also included a behavioral-consequence question: asking participants to allocate a fixed amount of bonus money between the two decision-makers. Like Study 1D, Study

1E was designed so that the Objective Decision Quality of the more-constrained decision-maker was actually higher than that of the less-constrained decision-maker.

7.1 | Method

Participants were 154 corporate managers recruited from Credamo, an online survey platform in China (85 females, $M_{age} = 33.60$). They read a case similar to the Abby versus Barb cases in Study 1A and Study 1B. (In Study 1A and Study 1B, we described the more-constrained decision-maker—Abby—before the less-constrained decision-maker—Barb. For generality, in Study 1E, we described the less-constrained decision-maker—Larry—before the more-constrained decision-maker—Matt). Specifically, participants read the following case (originally in Chinese):

> The CEO of a company recently posted recruitment ads to hire employees for two newly-created departments—L and M. A total of 160 people responded to the ads and applied. Half of them applied to department M and the other half applied to department L. Some applicants were high-quality and some were low-quality.
>
> Among those who applied to department L, 40 were high-quality and 40 were low-quality; among those who applied to department M, 20 were high-quality and 60 were low-quality.
>
> The CEO asked two assistants—Larry and Matt—to make hiring decisions (i.e., identifying and hiring high-quality applicants from the applicant pool). The CEO assigned Larry to make hiring decisions for department L and Matt to make hiring decisions for department M. The assignment was random and did not reflect the relative ability of the two assistants.
>
> Based on the needs of each department, the CEO required Larry to hire 40 applicants for department L, and Matt to hire 60 applicants for department M.

Before reading the remainder of the case, participants were required to answer seven comprehension questions:

> In total, how many people responded to the ads and applied?
> How many people applied to department L?
> How many of them were high-quality applicants?
> How many people applied to department M?
> How many of them were high-quality applicants?
> How many applicants did the CEO require Larry to hire?
> How many applicants did the CEO require Matt to hire?

For each question, participants were given five possible answers: 20, 40, 60, 80, and other. The correct answers for the seven questions were other, 80, 40, 80, 20, 40, and 60, respectively. Every participant must answer all questions correctly before they were allowed to read the rest of the case; if they had answered any question incorrectly, they had to re-answer it until they answered it correctly.

After having correctly answered all the comprehension questions, participants were shown the rest of the case:

> The two assistants followed the CEO's instructions and hired the required numbers of applicants for their respective departments.
>
> Among the applicants Larry hired, 32 were high-quality and 8 were low-quality. Among the applicants Matt hired, 17 were high-quality and 43 were low-quality.
>
> How would you rate the quality of Larry's decisions? How would you rate the quality of Matt's decisions?

Participants then judged the quality of each decision-maker's decisions by choosing one of six options, ranging from "*very low*" (1) to "*very high*" (6).

Finally, participants were asked a bonus-allocation question:

> If you were the CEO and you had a fixed amount of bonus money to give Larry and Matt, how would you divide the money between them? Choose the closest answer below:

- 100% to Larry and 0% to Matt
- 80% to Larry and 20% to Matt
- 60% to Larry and 40% to Matt
- 40% to Larry and 60% to Matt
- 20% to Larry and 80% to Matt
- 0% to Larry and 100% to Matt

7.2 | Results and Discussion

As Table 5 shows, Matt was more constrained than Larry both in terms of #InHQ (20 vs. 40) and #Needed (60 vs. 40). Despite the constraints, Matt hired 17 of the 20 high-quality candidates available for him to hire, while Larry hired only 32 of the 40 high-quality candidates available for him to hire. Therefore, Matt's Objective Decision Quality was higher than Larry's. Despite this, the participants again exhibited the Blaming-the-Strawless-Brickmaker effect; judging Matt's decision quality to be significantly lower Larry's, $t(150) = 11.59$, $p < 0.001$, Cohen's $d = 1.62$. Not only did the participants exhibit the Blaming-the-Strawless-Brickmaker effect in their judgments, they also displayed this bias in their bonus

TABLE 5 | Parameters and results of Study 1E.

	More-constrained decision-maker (Matt)	Less-constrained decision-maker (Larry)
#InHQ	20	40
#Needed	60	40
Best Possible Output Quality	20/60	40/40
Realized Output Quality	17/60	32/40
Objective Decision Quality	0.95	0.80
Judged Decision Quality (mean and SD)	2.89 (1.37)	4.77 (0.90)
Bonus (% given to each decision-maker)	36.9%	63.1%

Note: Although in this study we described the less-constrained decision-maker (Larry) before the more-constrained decision-maker (Matt), in the table we still present the more-constrained decision-maker to the left of the less-constrained decision-maker in order to be consistent with the format of the tables for the other studies.

allocation decision: They allocated significantly less bonus money to Matt than to Larry, 36.9% to Matt and 63.1% to Larry, $t(150) = 9.96$, $p < 0.001$, Cohen's $d = 0.81$, compared with equal allocation.

Study 1E demonstrated the robustness of the Blaming-the-Strawless-Brickmaker effect: It replicated the effect even when the participants were businesspeople and from a non-Western culture, even though the Objective Decision Quality of the more-constrained decision-maker was higher, even after the participants had passed all the comprehension questions, and even in their allocation of bonus money between the decision-makers. These results cast double on inattention and misunderstanding as alternative explanations.

8 | Study 1F

To further test the robustness and the behavioral consequence of the Blaming-the-Strawless-Brickmaker effect, we conducted Study 1F, an incentive-compatible study. In the study, the decision-makers were two ball-picking algorithms, their decisions were to pick balls from a bag, and high-quality candidates and low-quality candidates were winning balls and losing balls in the bag. Participants were informed of the results of a previous test, in which Alex faced more constraints but made better decisions. Participants were then asked whether they would delegate Alex or Bob to pick balls (i.e., make decisions) for them. The study entailed real financial consequences to the participants.

8.1 | Method

Participants were 150 online workers recruited from Prolific (73 females, $M_{age} = 30.63$). In the beginning, participants were told:

> By default, you will get a 25-cent bonus for doing this study. However, the actual amount of bonus you will get may be more or less than 25 cents. See below for details.

The participants were then shown the following picture and the ensuing instructions:

> Above is a virtual bag that contains 25 winning balls and 25 losing balls. We have two ball-picking algorithms, which we call Alex and Bob. In a moment, one of them (either Alex or Bob) will pick 25 balls from the bag for you. For each winning ball they pick, you will gain 1 cent in bonus money; for each losing ball they pick, you will lose 1 cent in bonus money. The more winning balls they pick, the more bonus money you will gain; the more losing balls they pick, the more bonus money you will lose.

> You can decide whether to have Alex or Bob pick balls for you. Alex and Bob have different tendencies to pick winning balls vs. losing balls.

> In a recent test, we asked Alex to pick 25 balls from a bag that contained 11 winning balls and 39 losing balls, and Bob to pick 25 balls from a bag that contained 24 winning balls and 26 losing balls.

> The results of the test were as follows: Alex picked 9 winning balls and 16 losing balls, and Bob picked 19 winning balls and 6 losing balls.

> As mentioned earlier, one of them (Alex or Bob) will pick 25 balls from the bag shown at the top. Do you want Alex or Bob to pick balls for you?

Participants then made their choice. At the end of the study, each participant received a \$0.41 bonus. The amount of bonus was calculated based on the decision quality of Alex and Bob (around 80%).

8.2 | Results and Discussion

According to the information summarized at the top of Table 6, in the previous test, Alex had only 11 winning-balls (high-quality candidates) in its pool, while Bob had 24 winning-balls in tis pool, so Alex was more constrained than Bob. But Alex's Objective Decision Quality was higher than Bob's.

Despite the difference, the majority of the participants (66.0%) chose to delegate Bob to pick balls for them, $\chi^2 = 15.36$, $p < 0.001$, compared with 50%. This result replicated the Blaming-the-Strawless-Brickmaker effect in a direct choice paradigm with real financial consequences to the participants. Because the dependent variable was choice (not judgment), the result further ruled out misunderstanding of the meaning of "decision quality" as an alternative explanation.

9 | Study 2

Now that the studies reported so far tested and supported our primary hypothesis (H1) regarding information usage (i.e., underweighting constraint-related information in the judgment phase), Study 2 tested our second hypothesis (H2) regarding information search (i.e., underweighting constrained-related information in the information acquisition phase).

Like Study 1F, Study 2 was also incentive-compatible. Participants were tasked to judge the relative decision quality of a more-constrained decision-maker (Abby) and a less-constrained decision-maker (Barb), and they would receive a bonus if their judgment was accurate.

Objectively speaking, in order to make the judgment accurately, the evaluator would need at least the following three pieces of information: (a) input quality (i.e., #InHQ among the candidates), (b) need (i.e., #Needed), and (c) output quality (i.e., Realized Output Quality). The first two pieces of information were about the constraints of the decision-makers, and the last piece of information was about their output quality.

The study consisted of two phases. In phase 1, the participants did not have any of the above information, but they could incur a monetary cost to acquire any or all of these pieces of information. The dependent variable in this phase was which piece(s) of information the participants would seek. In Phase 2, all participants were given all three pieces of information (even if they had not sought it in Phase 1), and were asked to judge the relative quality of Abby's and Barb's decisions. The dependent variable for Phase 2 was this judgment.

Based on H2, we predicted that before they had any of the information (i.e., in phase 1), participants were less likely to seek the two pieces of constraint-related information (#InHQ and #Needed) than to seek the output-quality information. Based on H1, we predicted that after participants had received all pieces of information (i.e., in Phase 2), they would still exhibit the Blaming-the-Strawless-Brickmaker effect, suggesting that they still underweighted the constraint-related information.

9.1 | Method

Participants in the study were 150 online workers recruited from Prolific (77 females, $M_{age} = 27.25$). In the beginning, they were told:

> By default, you will get a 15-cent bonus for completing this study, but the actual amount of bonus you will get may be more or less than 15 cents. See below for details.

The participants then entered Phase 1 of the study, which tested which type of information they would look for. The participants were

TABLE 6 | Parameters and results of Study 1F.

	More-constrained decision-maker (Alex)	Less-constrained decision-maker (Bob)
#InHQ	11	24
#Needed	25	25
Best Possible Output Quality	11/25	24/25
Realized Output Quality	9/25	19/25
Objective Decision Quality	0.92	0.80
Choice (% of participants)	34.0%	66.0%

> The CEO of a company recently posted recruitment ads to hire employees for two newly-created departments—A and B. A total of 100 people responded to the ads and applied. Half of them applied to department A and the other half applied to department B. Some applicants were high-quality, and the rest were low-quality.

> The CEO asked two assistants—Abby and Barb—to make hiring decisions (i.e., identifying and hiring high-quality applicants from the applicant pool). The CEO assigned Abby to make hiring decisions for department A and Barb to make hiring decisions for department B. The assignment was random and did not reflect the relative ability of the two assistants. Based on the demand of each department, the CEO required each assistant to hire a specific number of applicants. The two assistants followed the CEO's instructions.

> Your task is to judge the relative quality of Abby's and Barb's decisions. If your judgment is accurate, you will get an extra 40-cent bonus.

> Before you make your judgment, you have the option to use your current bonus money (15 cents) to get some additional pieces of information. The following table tells you which additional pieces of information you can get. In order to get the information on each row below, you will lose 5 cents of your current bonus money. It's up to you which additional piece(s) of information to get.

<box>	Among the applicants to each department (A vs. B), what proportion was high-quality
<box>	How many applicants each assistant (Abby vs. Barb) was required to hire
<box>	Among the applicants each assistant (Abby vs. Barb) eventually hired, what proportion was high-quality

> To judge the relative quality of Abby's and Barb's decisions, which additional piece(s) of information do you want to get? Check the corresponding box (es) above. Remember: for each box you check, you will lose 5 cents of your current bonus money.

After the participants checked the box(es), they were guided to a new screen, and entered phase 2 of the study, in which they were given all the three pieces of information and asked to judge the relative decision quality of the two decision-makers. The participants read:

> Regardless of which boxes checked on the previous page, we give you the following information:

> After the CEO had posted the recruitment ads, 50 people applied to department A, and 50 people applied to department B. (The two groups of people did not overlap.) Among those who applied to department A, 12 were high-quality and 38 were low-quality; among those who applied to department B, 24 were high-quality and 26 were low-quality.

> The CEO required Abby to hire 36 applicants for department A, and Barb to hire 23 applicants for department B. Abby and Barb followed the CEO's instructions and hired the required numbers of applicants for their respective departments.

> Among the applicants Abby eventually hired, 11 were high-quality and 25 were low-quality; among the applicants Barb eventually hired, 20 were high-quality and 3 were low-quality.

> How would you judge the relative quality of Abby's and Barb's decisions?

- Abby's decisions were much better than Barb's
- Abby's decisions were slightly better than Barb's
- Barb's decisions were slightly better than Abby's
- Barb's decisions were much better than Abby's

As we will explain below, Abby's Objective Decision Quality was higher than Barb's. Therefore, as long as a participant judged Abby's decisions as better than Barb's (regardless of whether the participant judged Abby's decisions as much better or slightly better than Barb's), we gave them a 40-cent bonus.

9.2 | Results and Discussion

9.2.1 | Phase 1 Results (Information Search)

In support of H2, significantly more participants sought the output-quality information (56.0%) than participants who sought either of the two pieces of constraint-related information: output quality versus input quality (38.0%), $\chi^2 = 6.56$, $p = 0.010$; output quality versus need (34.7%), $\chi^2 = 9.61$, $p = 0.002$, by McNemar test.

Because every participant could choose more than one piece of information, we also calculated the percentage of participants who chose different combinations of information, and report the results in Table 7.1. Again consistent with our theory, only a small minority (9.3%) of the participants sought all three pieces of information. The modal participant sought only the output-quality information, and the proportion of such participants (40.7%) was significantly greater than the proportion of participants who sought any other combinations of information, $\chi^2(2, 150) = 9.04$, $p = 0.011$, when compared with the next highest proportion.

TABLE 7.1 | Percentage of participants seeking different combinations of information in Study 2.

Combination of information	Percentage of participants
All three pieces of information	9.3%
Only output quality	40.7%
Only input quality	21.3%
Only need	18.7%
Input quality and output quality	3.3%
Input quality and need	4.0%
Need and output quality	2.7%
None of the three pieces of information	0%

TABLE 7.2 | Parameters and judgment results of Study 2.

	More-constrained decision-maker (Abby)	Less-constrained decision-maker (Barb)
#InHQ	12	24
#Needed	36	23
Best Possible Output Quality	12/36	23/23
Realized Output Quality	11/36	20/23
Objective Decision Quality	0.97	0.87
Judged Decision Quality (% of participants judging each decision-maker's decisions as better than the other's)	38.0%	62.0%

Note: The exact percentage of participants who chose each judgment option was as follows: 10.7% judging Abby's decisions as much better than Barb's, 27.3% judging Abby's decisions as slightly better than Barb's, 19.3% judging Barb's decisions as slightly better than Abby's, and 42.7% judging Barb's decisions as much better than Abby's.

9.2.2 | Phase 2 Results (Judged Decision Quality)

Table 7.2 summarizes the parameters and the results. As the table shows, Abby was more constrained than Barb, because Abby (vs. Barb) had fewer high-quality candidates in her candidate pool (12 vs. 24), yet needed to accept more candidates (36 vs. 23). Abby eventually accepted 11 of the 12 high-quality candidates available for her to accept, while Barb accepted only 20 of the 23 high-quality candidates she could have hired. Therefore, Abby's Objective Decision Quality was higher than Barb's.

However, as the bottom section of Table 7.2 presents, only a minority of the participants (38.0%) judged Abby's decisions as better than Barb's; the majority (62.0%) considered Barb's decisions as better, $\chi^2 = 8.64$, $p = 0.003$, compared with 50%. Consistent with H1, these results suggest that even after the evaluators had received all the relevant information, they still underweighted the constraint-related information.

In summary, Study 2 suggests that evaluators overlook constraint-related information both in the information search phase and in the judgment phase.

10 | Study 3

Study 3 tested our last hypothesis—H3, regarding the moderating effect of prompting the evaluator to consider the constraints of the decision-makers. Notably, the prompt merely nudged the evaluator to consider the constraints of the decision-maker; it did not provide any new information, or make any directional suggestions. If this prompt mitigated the Blaming-the-Strawless-Brickmaker effect (as predicted by H3), it would further support our constraint-neglect theory.

10.1 | Method

Participants in the study were 300 online workers recruited from Prolific (146 females, $M_{age} = 31.35$). They were randomly assigned to one of two conditions—control and treatment. All participants read the following case:

> The CEO of a company recently posted recruitment ads to hire employees for two newly-created departments—A and B. A total of 100 people responded to the ads and applied. Half of them applied to department A and the other half applied to department B. Among those who applied to department A, 12 were high-quality and 38 were low-quality; among those who applied to department B, 24 were high-quality and 26 were low-quality.

> The CEO asked two assistants—Abby and Barb—to make hiring decisions (i.e., identifying and hiring high-quality applicants from the applicant pool). The CEO assigned Abby to make hiring decisions for department A and Barb to make hiring decisions for department B. The assignment was random and did not reflect the relative ability of the two assistants.

> The CEO required Abby to hire 36 applicants for department A, and Barb to hire 23 applicants for department B. Abby and Barb followed the CEO's

> instructions and hired the required numbers of applicants for their respective departments.

> Among the applicants Abby hired, 11 were high-quality and 25 were low-quality; among the applicants Barb hired, 20 were high-quality and 3 were low-quality.

Participants in the control condition were then asked to judge the quality of each person's decisions. For each judgment, they were given six options (ranging from "extremely low" to "extremely high"), as in Study 1A and Study 1B.

Participants in the treatment condition were asked the same judgment questions, but were first given the following prompt:

> Before you answer, please consider the constraints Abby vs. Barb faced: how many high-quality applicants were available for Abby to hire vs. how many high-quality applicants were available for Barb to hire, and how many applicants Abby were required to hire vs. how many applicants Barb were required to hire. Given these constraints, how would you rate the quality of each person's decisions?

10.2 | Results and Discussion

As the top section of Table 8 displays, the stimuli in Study 3 were similar to the stimuli in Study 2, such that the Objective Decision Quality of the more-constrained decision-maker (Abby) was better than that of the less-constrained decision-maker (Barb).

The last two rows of Table 8 present the results. A one-way MANOVA found that the presence or absence of the prompt had a significantly different effect on judged decision quality, $F(2, 297) = 4.64$, $p = 0.01$. In the control condition, participants again displayed the Blaming-the-Strawless-Brickmaker effect, giving a significantly lower rating to Abby's decisions than to Barb's decisions, $p < 0.001$. In the treatment condition, the effect disappeared, though it did not reverse, $p = 0.487$. It seems that the prompt for evaluators to consider the decision-makers' constraints was powerful enough to eliminate the Blaming-the-Strawless-Brickmaker effect, but was not powerful enough to reverse it.

The results also reveal that the prompt had asymmetrical effects on the judgments of the two decision-makers: It significantly increased the judged quality of Abby (the more-constrained decision-maker): 3.85 versus 4.29, $p < 0.001$, and only marginally decreased the judged quality of Barb (the less-constrained decision-maker): 4.62 versus 4.40, $p = 0.062$. This asymmetry is consistent with the fact that the Objective Decision Quality of both Abby and Barb was very high (0.97 vs. 0.87): In the control condition, the Judged Decision Quality of Abby was unfairly low, while the Judged Decision Quality of Barb was already relatively high. The prompt in the treatment condition significantly improved the Judged Decision Quality of Abby fairer, and did not equally worsen the judgment of Barb, which was already relatively fair in the control condition.

The finding of this study lent further to our constraint-neglect theory; at the same time, it cast doubt on lack of mathematical skills as an alternative explanation for the Blaming-the-Strawless-Brickmaker effect. If the effect were due to a lack of mathematical skills, the prompt would not have had a significant moderating effect, because the prompt did not give the participants any suggestions about how to perform the math in the problem.

11 | General Discussion

While extensive existing research has studied actual decision quality, the present research studies the judgment of decision quality. We propose a constraint-neglect theory about the judgment. In support of the theory, eight studies provide strong evidence for the Blaming-the-Strawless-Brickmaker effect, a tendency to judge a decision-maker who encountered greater constraints as having made worse decisions than a decision-maker who encountered lesser constraints, although the former actually made as good decisions as or even better decisions than the latter. Our effect is robust to different contexts, different methods, and different types of participants. Our research also

TABLE 8 | Parameters and results of Study 3.

	More-constrained decision-maker (Abby)	Less-constrained decision-maker (Barb)
#InHQ	12	24
#Needed	36	23
Best Possible Output Quality	12/36	23/23
Realized Output Quality	11/36	20/23
Objective Decision Quality	0.97	0.87
Judged Decision Quality		
Control condition (mean and SD)	3.85 (1.50)	4.62 (1.03)
Treatment condition (mean and SD)	4.29 (1.27)	4.40 (1.01)

demonstrates the behavioral consequence of the Blaming-the-Strawless-Brickmaker effect, tests its underlying process, and suggests ways to mitigate the bias. In the rest of the article, we discuss open questions and implications.

11.1 | Do Evaluators Under-Appreciate the More-Constrained Decision-Maker or Over-Appreciate the Less-Constrained Decision-Maker?

Our studies demonstrate that the evaluator under-appreciates the more-constrained decision-maker *relative to* the less-constrained decision-maker, but do not specify whether the evaluator under-appreciates the more-constrained decision-maker in the absolute sense, over-appreciates the less-constrained decision-maker in the absolute, or both.

Although we cannot answer the above question for sure, we have two pieces of evidence in favor of the first possibility—under-appreciating the more-constrained decision-maker. First, in Study 1A and Study 1B, the Objective Decision Quality was 1.00 (i.e., the best possible level) for both the more-constrained decision-maker (Abby) and the less-constrained decision-maker (Barb), but the Judged Decision Quality for Abby was further from the best possible level of the judgment scale than the Judged Decision Quality for Barb. In other words, although both Abby and Barb did the best they could, the Judged Decision Quality for Abby was more harsh than that for Barb, which suggests that the participants under-appreciated Abby (the more-constrained decision-maker) rather than over-appreciated Barb (the less-constrained decision-maker).

Second, as noted earlier, the prompt manipulation in Study 3 had asymmetric effects: It increased the Judged Decision Quality of the more-constrained decision-maker (Abby) more than it decreased the Judged Decision Quality of the less-constrained decision-maker (Barb). To the extent that the prompt manipulation made participants' judgments more accurate, the asymmetric effects are evidence that in the control condition, the judgment for the more-constrained decision-maker was too low, rather than the judgment for the less-constrained decision-maker was too high.

These data support the possibility that the evaluator under-appreciates the more-constrained decision-maker rather than over-appreciates the less-constrained decision-maker, hence justifying the term "Blaming the Strawless Brickmaker" (instead of Praising the Strawful Brickmaker). This pattern also agrees with previous literature showing that people are more likely to attribute others' negative behavior than others' positive behavior to dispositional factors (Malle 2006). Nevertheless, further research is needed to ascertain whether this asymmetry is general, or specific to the stimuli of our studies.

11.2 | What If the Evaluator Lacks Information About the Constraints?

This research focuses on situations in which the evaluator has information about the constraints of the decision-maker—knowing how many high-quality versus low-quality candidates there were in the candidate pool, and how many candidates the decision-maker needed to accept. As alluded to in the Introduction, evaluators in real life often lack such information. Our theory predicts—and the results of Study 2 show—that when the evaluator does not have such information, they are unlikely to expend resources to seek it. Evaluators in such real-life situations, relative to participants in our experiments, are even more likely to base their judgments solely on the output quality of the decision-maker. Thus, we believe that the Blaming-the-Strawless Brickmaker effect is more prevalent in real life than in our experiments.

11.3 | What If Quality Is a Continuous Variable?

It is for ease of analysis that we assume that the quality of a candidate is dichotomous—either high or low. Our theory still applies if quality is a continuous variable. Suppose that the quality of a candidate can range from 0 (*lowest*) to 1 (*highest*). Then the Best Possible Output Quality is simply the average quality of the #Needed highest-quality candidates in the candidate pool, the Realized Output Quality is simply the average quality of the candidates the decision-maker eventually accepts, and the Objective Decision Quality is the ratio of the Realized Output Quality over the Best Possible Output Quality.

For example, suppose that decision-maker Abby needs to accept 5 candidates from a pool of 8 candidates, whose qualities are 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, and 0.1, respectively, and decision-maker Barb needs to accept 3 candidates from a pool of eight candidates, whose qualities are 1.0, 1.0, 1.0, 0.9, 0.9, 0.8, 0.8, and 0.8, respectively. Suppose also that Abby eventually accepts the 5 highest-quality candidates in her pool, while Barb eventually accepts the 3 lowest-quality candidates in her pool. Then for Abby, the Best Possible Output Quality is the average of 0.8, 0.7, 0.6, 0.5, and 0.4, namely, is 0.6, her Realized Output Quality is also 0.6, and so her Objective Decision Quality is 1.0. For Barb, the Best Possible Output Quality is 1.0, but her Realized Output Quality is only 0.8, and so her Objective Decision Quality is 0.8. Thus, Abby's Objective Decision Quality is higher than Barb's.

Based on our theory that Judged Decision Quality depends largely on Realized Output Quality, we predict that Abby will be judged as having made worse decisions than Barb, because Abby's Realized Output Quality is only 0.6 while Barb's Realized Output Quality is 0.8. This is an example of the Blaming-the-Strawless-Brickmaker effect in situations where the quality of candidates is a continuous variable.

In short, when the quality of candidates is a continuous variable,

$$\begin{aligned}&\text{Best Possible Output Quality}\\&=\text{average quality of the top\#Needed candidates}\end{aligned} \tag{5}$$

$$\text{Realized Output Quality} = \text{average quality of the accepted candidates} \tag{6}$$

Notably, Equations (5) and (6) are generalized versions of the corresponding equations for the dichotomous-quality case (Equations (1) and (2)). If we code the quality of a low-quality

candidate as 0, and the quality of a high-quality candidate as 1, then Equation (5) will yield the same value as Equation (1), and Equation (6) the same value as Equation (2).

11.4 | What Other Factors Unbiased Evaluators Should Consider?

In the Introduction, we posited that unbiased evaluators should base their judgment of one's decision quality on their Objective Decision Quality. Besides Objective Decision Quality, there are other factors that unbiased evaluator may need to consider. One is how likely the outcome of a decision-maker is due to mere luck. Ceteris paribus, the less likely one's outcome is due to mere luck, the more likely it is due to their decision quality. We already discussed this factor in Study 1C and in Footnote 3, and ruled out this factor as a viable alternative explanation for our result.

Another factor that unbiased evaluators may need to consider is how difficult it is to distinguish high-quality candidates from low-quality candidates in a given pool. Ceteris paribus, the more difficult it is to distinguish high-quality candidates from low-quality candidates, the better one's decisions. Although this is a factor that unbiased evaluations may need to consider, it does not explain the results of the present research, because there is no reason to believe that the high-quality and the low-quality candidates were systematically more distinguish or less distinguishable in the more-constrained decision-maker's pool versus in the less-constrained decision-maker's pool.

In short, although there are other potential factors than Objective Decision Quality that unbiased (rational) evaluators *should* consider, these factors are orthogonal to the theory and the findings of the present research, which is primarily about what biased (regular) evaluators *actually* do.

11.5 | Implications for Judgment of Discrimination

Although this research concerns the judgment of decision quality, it has implications for the judgment of discrimination (Hsee and Li 2022; Phillips and Jun 2022; Schmitt et al. 2014). Suppose that two new departments of a company (F and M) are each recruiting new employees. Most of the applicants to department F are female, most of the applicants to department M are male, and, on average, the female and the male applicants are equally qualified. Two managers—Feng and Meng—are assigned to make hiring decisions for departments F and M, respectively. Then, even if both managers are gender-blind, the unequal compositions of females and males in their respective candidate pools will lead Feng to hire more females than males, and Meng to hire more males than females. In this case, we predict that the evaluator will judge Feng as discriminating against males, and Meng as discriminating against females, because, according to the constraint-neglect theory, the evaluator tends to base their decision on the output information (i.e., the proportion of females vs. males among the hired candidates) and overlook the input information (i.e., the proportion of females versus males in the candidate pool).

11.6 | Practical Implications

This research provides practical insights for the evaluators in an organization, the decision-maker in the organization, and the organization itself. Regarding the evaluator, our data suggests that when they do not have constraint-related information, they do not actively seek the information, and when they have the information, they do not sufficiently use it unless they are prompted to consider it. These findings suggest two ways for evaluators to make more accurate judgments. First, they should realize that constraint-related information is important but is often unavailable, and they should actively seek such information in order to make accurate judgments. Second, once they have the constraint-related information, they should deliberately and correctly use it; specifically, they should consider what the best possible outcome a decision-maker could have achieved given the constraints she faced, and how far her actual decisions fall short of this best possible outcome. If possible, the evaluator should figure out the decision-maker's Best Possible Output Quality *before* finding out their Realized Output Quality; otherwise, the analysis of the Best Possible Output Quality may be anchored on and biased by the Realized Output Quality (Tversky and Kahneman 1974) and insufficient adjustment to the fair evaluation (Epley and Gilovich 2006).

Regarding the decision-maker, our research suggests that they should be aware of and foresee the bias the evaluator may have when judging their decision quality. If the decision-maker has encountered great constraints, they should actively inform their evaluator (e.g., their superiors) of the constraints, and nudge them to correctly incorporate the information in their judgment (or by reading our paper).

The current research also yields practical implications at the organizational level. The Blaming-the-Strawless-Brickmaker effect is detrimental to organizations. If the management of an organization under-appreciates "skillful but strawless brickmakers"—decision-makers who have done their best given the constraints they encountered, these people may feel mistreated and leave the organization, or they may strategically avoid making difficult decisions (decisions with greater constraints), and look for easy decisions to make. To combat the Blaming-the-Strawless-Brickmaker problem, organizations should have a formal decision-quality evaluation system that explicitly incorporates the constraints of the decision-maker (like Equation (3) or (5)).

Furthermore, in order to keep the overall quality of the accepted candidates high, organizations should strive to improve not only the decision quality, but also the input quality, namely, to attract the best possible candidates in the first place. If the input quality is inevitably low, and if the need is flexible, organizations may want to lower the input quantity. The rationale behind these two strategies is to optimize the two constraint variables, respectively, namely, to improve #InHQ, and limit #Needed. In the context of a journal, these strategies mean that the editor should work to attract the best possible submissions (rather than just try to accept the best submitted manuscripts or help improve the submitted manuscripts), and, if there are not enough good submissions and the length of the journal is flexible, the editor may want to reduce the number of accepted manuscripts so as to keep the overall quality of the journal high. Likewise, the HR office of a firm and the admissions office of a college may also

want to adopt these strategies in order to keep the overall quality of the institution high.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Endnotes

[1] Another way to express Equation (1) is

$$\textit{Best Possible Output Quality} = 1 \;\; if \# InHQ > \# Needed$$

$$\text{or} = \frac{\# InHQ}{\# Needed} \;\; if \# InHQ \leq \# Needed$$

This expression is mathematically identical to Equation (1) in the main text.

[2] Another way to define Objective Decision Quality is to use the ratio (instead of the difference) between Realized Output Quality and Best Possible Output Quality. We are agnostic whether it is better to use ratio or difference, but, regardless of whether to use ratio or difference, the rationale is the same: The judgment of Objective Decision Quality should use Best Possible Output Quality as the benchmark.

[3] Suppose that a candidate pool contains #InTotal candidates, of whom, #InHQ are high-quality and #InLQ are low-quality. A decision-maker is required to pick #OutTotal candidates. Of the #OutTotal candidates picked by the decision-maker, #OutHQ are high-quality and #OutLQ are low-quality. Then, statistically, the likelihood that this outcome is achieved merely by chance (luck) is

$$\frac{(\# InHQ)C(\# OutHQ) \times (\# InLQ)C(\# OutLQ)}{(\# InTotal)C(\# OutTotal)}$$

Based on this formula, the likelihood that Alex achieved what he achieved merely by chance is only 0.7%, while the likelihood that Bob achieved what he achieved merely by chance is almost 10 times as large—6.7%. By this logic, one could argue that Alex must have made better decisions than Bob.

References

Arkes, H. R. 1991. "Costs and Benefits of Judgment Errors: Implications for Debiasing." *Psychological Bulletin* 110, no. 3: 486–498.

Bar-Hillel, M. 1980. "The Base-Rate Fallacy in Probability Judgments." *Acta Psychologica* 44, no. 3: 211–233.

Baron, J., and J. C. Hershey. 1988. "Outcome Bias in Decision Evaluation." *Journal of Personality and Social Psychology* 54, no. 4: 569–579.

Bazerman, M. H., G. F. Loewenstein, and S. B. White. 1992. "Reversals of Preference in Allocation Decisions: Judging an Alternative Versus Choosing Among Alternatives." *Administrative Science Quarterly* 37, no. 2: 220–240.

Brogden, H. E., and E. K. Taylor. 1950. "The Theory and Classification of Criterion Bias." *Educational and Psychological Measurement* 10, no. 2: 159–183.

Dane, E., K. W. Rockmann, and M. G. Pratt. 2012. "When Should I Trust my Gut? Linking Domain Expertise to Intuitive Decision-Making Effectiveness." *Organizational Behavior and Human Decision Processes* 119, no. 2: 187–194.

Derfler-Rozin, R., and M. Pitesa. 2020. "Motivation Purity Bias: Expression of Extrinsic Motivation Undermines Perceived Intrinsic Motivation and Engenders bias in Selection Decisions." *Academy of Management Journal* 63, no. 6: 1840–1864.

Dooley, R. S., and G. E. Fryxell. 1999. "Attaining Decision Quality and Commitment From Dissent: The Moderating Effects of Loyalty and Competence in Strategic Decision-Making Teams." *Academy of Management Journal* 42, no. 4: 389–402.

Epley, N., and T. Gilovich. 2006. "The Anchoring-and-Adjustment Heuristic: Why the Adjustments Are Insufficient." *Psychological Science* 17, no. 4: 311–318.

Fisher, C. M. 2017. "An Ounce of Prevention or a Pound of Cure? Two Experiments on in-Process Interventions in Decision-Making Groups." *Organizational Behavior and Human Decision Processes* 138: 59–73.

Frederick, S., and B. Fischhoff. 1998. "Scope (In)sensitivity in Elicited Valuations." *Risk Decision and Policy* 3, no. 2: 109–123.

Frederick, S., N. Novemsky, J. Wang, R. Dhar, and S. Nowlis. 2009. "Opportunity Cost Neglect." *Journal of Consumer Research* 36, no. 4: 553–561.

Fredrickson, B. L., and D. Kahneman. 1993. "Duration Neglect in Retrospective Evaluations of Affective Episodes." *Journal of Personality and Social Psychology* 65, no. 1: 45–55.

Gilbert, D. T., and P. S. Malone. 1995. "The Correspondence Bias." *Psychological Bulletin* 117, no. 1: 21–38.

Hsee, C. K. 1996. "The Evaluability Hypothesis: An Explanation for Preference Reversals Between Joint and Separate Evaluations of Alternatives." *Organizational Behavior and Human Decision Processes* 67, no. 3: 247–257.

Hsee, C. K., and X. Li. 2022. "A Framing Effect in the Judgment of Discrimination." *Proceedings of the National Academy of Sciences* 119, no. 47: e2205988119.

Hsee, C. K., and Y. Rottenstreich. 2004. "Music, Pandas, and Muggers: On the Affective Psychology of Value." *Journal of Experimental Psychology: General* 133, no. 1: 23–30.

Hsee, C. K., and J. Zhang. 2010. "General Evaluability Theory." *Perspectives on Psychological Science* 5, no. 4: 343–355.

Kahneman, D. 2011. *Thinking, Fast and Slow*. New York, NY: Farrar, Straus and Giroux.

Kahneman, D., and A. Tversky. 1973. "On the Psychology of Prediction." *Psychological Review* 80, no. 4: 237–251.

Leclerc, F., C. K. Hsee, and J. C. Nunes. 2005. "Narrow Focusing: Why the Relative Position of a Good in Its Category Matters More Than It Should." *Marketing Science* 24, no. 2: 194–205.

Lefgren, L., B. Platt, and J. Price. 2015. "Sticking With What (Barely) Worked: A Test of Outcome bias." *Management Science* 61, no. 5: 1121–1136.

Lerner, J. S., Y. Li, P. Valdesolo, and K. S. Kassam. 2015. "Emotion and Decision Making." *Annual Review of Psychology* 66, no. 1: 799–823.

Li, X., and C. K. Hsee. 2019a. "Being "Rational" Is Not Always Rational: Encouraging People to Be Rational Leads to Hedonically Suboptimal Decisions." *Journal of the Association for Consumer Research* 4, no. 2: 115–124.

Li, X., and C. K. Hsee. 2019b. "Beyond Preference Reversal: Distinguishing Justifiability From Evaluability in Joint Versus Single Evaluations." *Organizational Behavior and Human Decision Processes* 153: 63–74.

Li, X., and C. K. Hsee. 2021a. "Free-Riding and Cost-Bearing in Discrimination." *Organizational Behavior and Human Decision Processes* 163: 80–90.

Li, X., and C. K. Hsee. 2021b. "The Psychology of Marginal Utility." *Journal of Consumer Research* 48, no. 1: 169–188.

Liersch, M. J., and C. R. M. McKenzie. 2009. "Duration Neglect by Numbers—And Its Elimination by Graphs." *Organizational Behavior and Human Decision Processes* 108, no. 2: 303–314.

Lurie, N. H., and J. M. Swaminathan. 2009. "Is Timely Information Always Better? The Effect of Feedback Frequency on Decision Making." *Organizational Behavior and Human Decision Processes* 108, no. 2: 315–329.

Malle, B. F. 2006. "The Actor-Observer Asymmetry in Attribution: A (Surprising) Meta-Analysis." *Psychological Bulletin* 132, no. 6: 895–919.

Moore, D. A., S. A. Swift, Z. S. Sharek, and F. Gino. 2010. "Correspondence bias in Performance Evaluation: Why Grade Inflation Works." *Personality and Social Psychology Bulletin* 36, no. 6: 843–852.

O'Reilly, C. A. 1982. "Variations in Decision makers' use of Information Sources: The Impact of Quality and Accessibility of Information." *Academy of Management Journal* 25, no. 4: 756–771.

Phillips, L. T., and S. Jun. 2022. "Why Benefiting From Discrimination Is Less Recognized as Discrimination." *Journal of Personality and Social Psychology* 122, no. 5: 825–852.

Read, D., G. Loewenstein, and M. Rabin. 1999. "Choice Bracketing." *Journal of Risk and Uncertainty* 19, no. 1–3: 171–197.

Ross, L. 1977. "The Intuitive Psychologist and His Shortcomings: Distortions in the Attribution Process." *Advances in Experimental Social Psychology* 10: 173–220.

Savani, K., and D. King. 2015. "Perceiving Outcomes as Determined by External Forces: The Role of Event Construal in Attenuating the Outcome bias." *Organizational Behavior and Human Decision Processes* 130: 136–146.

Schmitt, M. T., N. R. Branscombe, T. Postmes, and A. Garcia. 2014. "The Consequences of Perceived Discrimination for Psychological Well-Being: A meta-Analytic Review." *Psychological Bulletin* 140, no. 4: 921–948.

Scopelliti, I., H. L. Min, E. McCormick, K. S. Kassam, and C. K. Morewedge. 2018. "Individual Differences in Correspondence Bias: Measurement, Consequences, and Correction of Biased Interpersonal Attributions." *Management Science* 64, no. 4: 1879–1910.

Sellier, A. L., I. Scopelliti, and C. K. Morewedge. 2019. "Debiasing Training Improves Decision Making in the Field." *Psychological Science* 30, no. 9: 1371–1379.

Seo, M. G., and L. F. Barrett. 2007. "Being Emotional During Decision Making—Good or Bad? An Empirical Investigation." *Academy of Management Journal* 50, no. 4: 923–940.

Taylor, H. C., and J. T. Russell. 1939. "The Relationship of Validity Coefficients to the Practical Effectiveness of Tests in Selection: Discussion and Tables." *Journal of Applied Psychology* 23, no. 5: 565–578.

Tversky, A., and D. Kahneman. 1974. "Judgment Under Uncertainty: Heuristics and Biases." *Science* 185, no. 4157: 1124–1131.

Welsh, M. B., and D. J. Navarro. 2012. "Seeing Is Believing: Priors, Trust, and Base Rate Neglect." *Organizational Behavior and Human Decision Processes* 119, no. 1: 1–14.

Yoon, H., I. Scopelliti, and C. K. Morewedge. 2021. "Decision Making can Be Improved Through Observational Learning." *Organizational Behavior and Human Decision Processes* 137: 13–26.