Available online at www.sciencedirect.com
SCIENCE DIRECT®
ACADEMIC PRESS
Organizational Behavior and Human Decision Processes 91 (2003) 243–253
ORGANIZATIONAL BEHAVIOR AND HUMAN DECISION PROCESSES
www.elsevier.com/locate/obhdp

The input bias: The misuse of input information in judgments of outcomes

Karen R. Chinander[a,*] and Maurice E. Schweitzer[b]

[a] Department of Management, School of Business Administration, University of Miami, 414 Jenkins Building, Coral Gables, FL 33124, USA
[b] Operations and Information Management Department, Wharton School, University of Pennsylvania, 566 JMHH, Philadelphia, PA 19104, USA

Abstract

In this paper we identify an input bias, the systematic misuse of input information in judgments of outcome quality. In many settings irrelevant input measures, such as the amount of time an employee spends in the office, influence outcome assessments, such as performance reviews. Across four studies we find that input values subtly, but significantly distort judgments of outcome quality. Irrelevant input information predictably influences outcome assessments even when people recognize that input measures should not matter and believe that input information did not matter. We examine the mechanics of the input bias, and suggest that because input measures are often easy to manipulate or misrepresent, the input bias is likely to have broad implications for managerial judgment and decision making.

Keywords: Input bias; Outcome assessment; Managerial judgment; Performance reviews; Quality

Introduction

Judgments of quality are essential prerequisites for many decision making tasks. For example, prior to making a hiring decision a manager needs to assess the quality of alternative candidates. Similarly, prior to making a purchase decision, a consumer needs to assess the quality of alternative commodities. Despite the importance of this judgment process and the frequency with which people make these judgments, quality assessments are often very difficult to make. For instance, subjective outcomes are difficult to value and comparative standards are often difficult to construct (Hsee, Blount, Loewenstein, & Bazerman, 1999).

This paper explores the influence of irrelevant input information on judgments of outcome quality. Our results suggest that people automatically associate input quantity information with output quality, and we find that input information influences assessments of outcomes even when people recognize that input information should be irrelevant. We report results from four studies that describe the relationship between input quantity (the amount of inputs used in producing an outcome, such as preparation time or production costs) and perceptions of outcome quality.

In many domains decision makers use input measures as a proxy for outcome measures. Inputs are often positively correlated with outcomes, and in many cases input measures are easier to assess than outcomes. For example, a substantial stream of work uses research and development expenditures, an input quantity, as a measure of the innovativeness of a firm, an outcome (Baysinger, Kosnik, & Turk, 1991; Graves, 1988; Hansen & Hill, 1991; Levin, Cohen, & Mowery, 1985; Schoenecker, Daellenbach, & McCarthy, 1995).

In general, the use of input measures as a proxy for outcome quality is appropriate when the relationship between inputs and outputs is direct, consistent, and unbiased. In many cases, however, these conditions do not hold. First, the relationship between inputs and outcomes is not always positive and monotonic. For example, longer hospital stays are not always better. In fact, longer hospital stays lead to a greater risk of nosocomial infection (Scott, 1997). Second, the relationship

* Corresponding author. Present address: Florida Atlantic University, John D. MacArthur Campus, 5353 Parkside Drive, Jupiter, FL 33458. Tel.: 1-561-799-8608; Fax: 1-561-799-8502.
E-mail addresses: kchinand@fau.edu (K.R. Chinander), schweitzer@wharton.upenn.edu (M.E. Schweitzer).

0749-5978/03/$ - see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S0749-5978(03)00025-6

between inputs and outcomes may be inconsistent across individuals and organizations. There are several reasons for why this might be the case. Some individuals or organizations may simply be more efficient than others. Alternatively, inputs may be measured differently across organizations. For example, firms may use different accounting rules for assessing their research expenditures. Third, strategic agents can often purposefully manipulate input measures (and even the perception of input measures) to bias outcome assessment. For example, an employee may spend long hours in the office to give the perception of high inputs. This issue of managing impressions can even influence the way organizations are structured. For example, the National Bicycle Industrial Company manufactures and delivers custom ordered bicycles three *weeks* after receiving an order, even though it takes the firm less than three *hours* to manufacture and assemble a bicycle (Moffat, 1990). The company could invest resources to speed delivery, but consumers may not appropriately value faster delivery. In fact, we conjecture that the slow delivery time gives customers the sense that their customized product, which purportedly took a long time to produce, has greater value.

Our investigation is related to prior work that demonstrates that normatively irrelevant information can influence judgment and decision processes (March, 1987). In many cases, people have a difficult time separating relevant from irrelevant information. For example, Camerer, Loewenstein, and Weber (1989) demonstrate that people are often unable to disentangle the amount of information they know from the amount of information they think others know. In other cases people significantly over-weight proxy attributes (Fischer, Damodaran, Laskey, & Lincoln, 1987) and even obviously irrelevant information (Silvera, Josephs, & Giesler, 2001). For example, Silvera et al. (2001) find that the number of practice problems provided can influence the effort participants expend preparing for a test. Similarly, Nisbett, Zukier, and Lemley (1981) demonstrate that irrelevant information can influence predictions. In Nisbett, Zukier, and Lemley's studies they find that when predicting a future outcome (e.g., how high an incoming college freshman's GPA will be), people decrease their reliance upon diagnostic cues (e.g., high school GPA) when provided with non-diagnostic information (e.g., the number of plants a student has). In related work, Bastardi and Shafir (1998) demonstrate that when decision makers lack irrelevant information they often delay making a decision until they receive this information, and then subsequently use this information in making their decision.

The influence of irrelevant information has also been demonstrated in a number of settings with direct economic implications. For example, even though the source of payments should not influence how money is spent, people construct mental accounts (Thaler, 1985) and spend money in ways that violate standard economic theory. Similarly, decision makers attend to prior "sunk" costs even though only future costs and benefits should influence their decisions (Arkes & Blumer, 1985; Bazerman, Giuliano, & Appelman, 1984; Garland, 1990; Staw, 1976). Irrelevant cost information also influences willingness to pay judgments. Baron and Maxwell (1996) demonstrate that cost information influences willingness to pay for services such as crime prevention even when participants believe costs are unrelated to the benefits.

Prior work has also examined the role of irrelevant information in judgments of outcomes. For example, Josephs, Giesler, and Silvera (1994) found that participants who had completed the same amount of work were more satisfied with their accomplishments when they had created a high stack of output (e.g., their pages of work were stapled to journals) than when they had created a short stack of output (e.g., their pages of work were simply single pages).

Prior work also demonstrates that people have difficulty assessing the relevance of diagnostic cues when judging outcomes (Baron & Hershey, 1988; Hawkins & Hastie, 1990). Specifically, when evaluating a prior decision people over-weight outcome information. Although both good and bad decisions can lead to successful and unsuccessful outcomes, judges assume that the correlation between good decisions and good outcomes (and bad decisions and bad outcomes) is much higher than it actually is. Baron and Hershey (1988) define this effect as the outcome bias.

In this paper we identify an *input bias*, the use of irrelevant input information in judgments of outcome quality. Although the quality of an outcome is often positively related to the quantity of the inputs used to produce that outcome, the relationship between input quantity and output quality is not automatic. In many cases inputs are misused, misrepresented, or even negatively related to outcome quality.

We also explore the psychological mechanism underlying the input bias. We conceptualize the input bias as an association based bias (Arkes, 1991) whereby people *automatically* associate high input quantities with high outcome quality, because input quantity and outcome quality are often positively correlated. We argue that the input bias represents a particularly important judgment bias, because strategic agents can inappropriately cue this association and manipulate or misrepresent input information.

Study 1

Our first study establishes a link between input information and judgments of quality. In this study participants rated two student presentations under one of two input conditions. We also investigate the salience of the relationship between input information and judgments of quality. We asked participants whether or not they recalled the input values, and whether or not they thought input information had influenced and should influence their judgments of quality.

Method

First, we recruited two participants to prepare and deliver two independent presentations about an emerging technology. The first presentation described electronic ink, and was 7.5 min long. The second presentation described optical switches, and was 13.5 min long. We recorded both presentations using standard VHS technology.

We then recruited 83 participants to perform a rating task as part of an optional class exercise. We randomly assigned participants to one of two between-subject conditions. All participants rated both presentations, and viewed the presentations in the same order. Participants rated the electronic ink presentation first and the optical switches presentation second.

Before viewing each presentation, participants received information sheets about the upcoming presentation. This sheet included the amount of time each person had spent preparing his or her presentation. This value was manipulated across conditions. In the first condition, raters were told that the electronic ink presenter spent 8 h and 34 min (a long preparation time) preparing for the presentation, and that the optical switches presenter spent 37 min (a short preparation time) preparing for the presentation. We refer to this as the *Decreasing* condition. In the second condition raters were told that the electronic ink presenter spent 37 min preparing for the presentation, and that the optical switches presenter spent 8 h and 34 min preparing for the presentation. We refer to this as the *Increasing* condition.

After viewing each presentation participants rated the presentation along 5 items: (1) the overall quality of the presentation, (2) the quality of the information presented, (3) the presenter's knowledge of the subject, (4) the organization of the presentation, and (5) the quality of the presenter's presentation skills. Each question used a 10-point response scale (1 = Poor, 10 = Excellent).

The scale reliability was high. For the electronic ink presentation with short and long preparation times, the Cronbach's α coefficients were 0.93 and 0.96, respectively. For the optical switches presentation with short and long preparation times the Cronbach's α coefficients were 0.91 and 0.92, respectively. As a result, we use a composite measure (the average) of the 5 items to represent participants' quality ratings in subsequent analysis.

After participants had completed their ratings, they responded to a final set of questions. These questions asked participants whether or not they remembered the time of preparation for each presentation, if the amount of preparation time had influenced their ratings for each presentation, and if they believed that the amount of preparation time should influence their ratings.

Results

A total of 83 participants completed the study. Forty-one participants were assigned to the decreasing condition and 42 were assigned to the increasing condition.[1] On average, participants were 21.6 years old, and just over half of the participants were male (53.0%).

Supporting our thesis, the preparation time manipulation significantly influenced quality assessments. As depicted in Figs. 1 and 2, participants exposed to the high input condition (long preparation time) rated the quality of the same presentation higher than participants exposed to the low input condition (short preparation time). The differences in the composite quality measures were significant for both the electronic ink presentation, 7.11 (1.95) versus 5.86 (1.57), $t(80) = 3.19$, $p < .01$, and the optical switches presentation, 7.38 (1.56) versus 6.62 (1.77), $t(80) = 2.06$, $p < .05$.

Overall, we found few differences in ratings between participants who thought input information should or should not matter and participants who thought that input information did or did not matter. In an analysis of variance model of quality judgments for the electronic ink presentation input time was significant, $F(1,76) = 9.65$, $p < .001$, but should not matter and did not matter parameters were not significant $F(2,76) = 0.57$, $p = n.s.$ and $F(2,76) = 0.91$, $p = n.s.$ We find similar results in an analysis of variance model of quality judgments for the optical switches presentation; input time was significant, $F(1,76) = 5.62$, $p < .05$, but should not matter and did not matter parameters were not significant $F(2,76) = 0.32$, $p = n.s.$ and $F(2,76) = 1.32$, $p = n.s.$

We next consider a more conservative set of tests regarding participants who thought input information should not and did not matter. Most participants, 63 (77%), thought input time should not influence their quality ratings. (Only 9 (11%) thought input time should influence their judgments, and 10 (12%) were not sure.) Even participants who believed input time should not matter, however, exhibited the same pattern of

[1] In the increasing group, one participant had a missing response for one of the electronic ink presentation questions, and another participant had a missing response for one of the optical switches presentation questions. One participant in the decreasing group did not respond to whether input time should matter. As a result, some analyses have missing values and lower degrees of freedom.

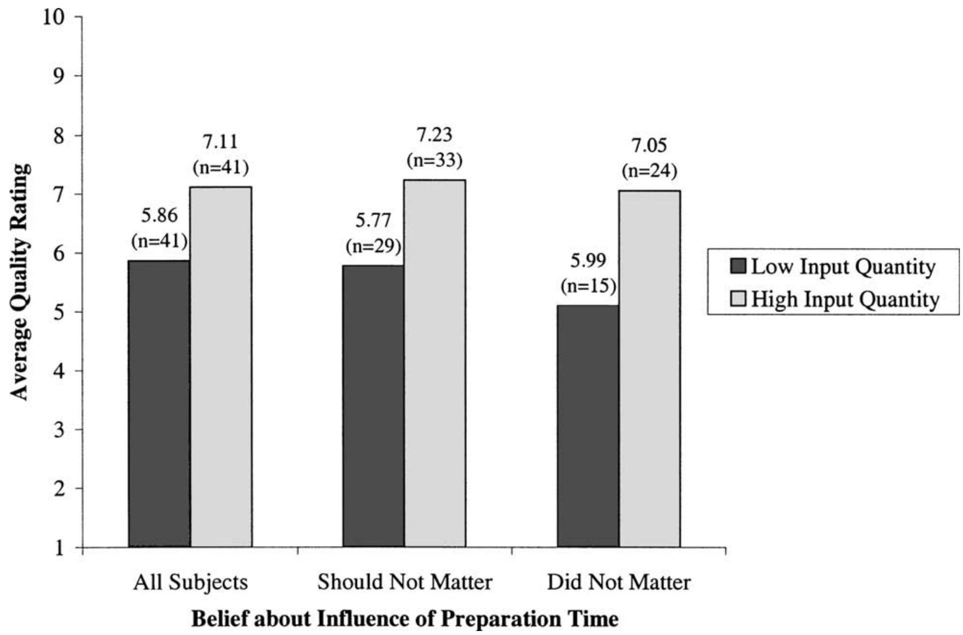

Fig. 1. Study 1—Electronic ink presentation ratings by input quantity.

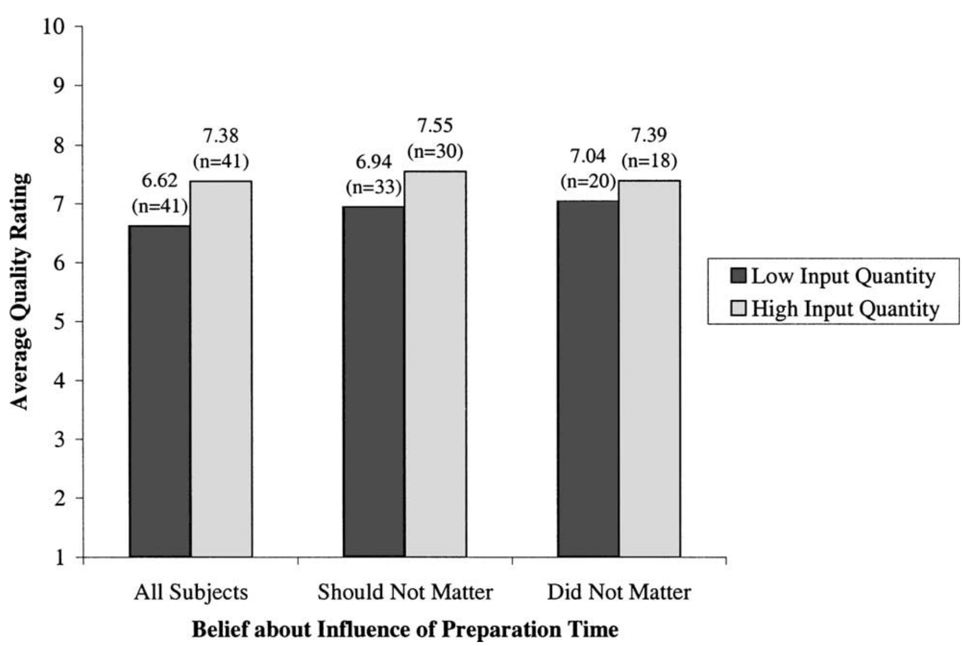

Fig. 2. Study 1—Optical switches presentation ratings by input quantity.

evaluations. Among these participants, those who were exposed to the high input condition rated the quality of the presentation higher than participants exposed to the low input condition for both presentations. This difference was significant for the electronic ink presentation, 7.23 (1.84) versus 5.77 (1.63), $t(60) = 3.29$, $p < .01$, but not significant for the optical switches presentation, 7.55 (1.36) versus 6.94 (1.75), $t(61) = 1.53$, $p = n.s.$

For each presentation participants also responded to questions about whether or not input time did influence their ratings. Consistent with previous results, even participants who believed that input information had not influenced their judgment rated each presentation higher when inputs (preparation time) were greater. For the electronic ink presentation the average ratings were 7.05 (2.07) versus 5.99 (2.00), and for the optical switches presentation average ratings were 7.39 (1.73) versus 7.04 (1.56); these ratings, however, are not significantly different, $t(37) = 1.58$, $p = n.s.$, and $t(36) = 0.65$, $p = n.s.$, respectively.

Discussion

Results from this study demonstrate that outcome assessments are influenced by input information. Participants in this study judged outcomes more favorably when preparation time was reported to be high than when preparation time was reported to be low. We find this same pattern of results even among participants who believe input time should not influence judgment.

Study 2

We extend our investigation of the input bias by considering an alternative set of inputs and outputs. In this study we gave participants information about the cost of machinery used to manufacture two food items, and we asked participants to taste and rate the relative quality of samples of these two foods.

Method

Participants in this study were recruited from the entrance to a college dormitory to participate in a taste test. Participants were randomly assigned to one of two input conditions. Prior to tasting and rating two different types of chocolate fudge, participants were provided with information about the machinery used to manufacture the fudge. The input conditions described either Fudge A or Fudge B as having been made with an expensive machine. Specifically, participants were informed before their taste test that both samples of fudge were "made with similar ingredients, however the ingredients were mixed and cooked differently. Fudge A was made using [very expensive/inexpensive] machinery, and Fudge B was made using [inexpensive/very expensive] machinery."

We use fudge in this study because the outcome assessment of fudge is direct and relatively simple. That is, the motivation for consuming fudge and the quality assessment of fudge is generally based upon taste. There are rarely other reasons (e.g., nutritional value) for consuming fudge.

After tasting both samples of fudge participants evaluated the relative quality of the samples on a 7-point scale (1: "Fudge A is much higher quality"; 7: "Fudge B is much higher quality"). After responding to this question, participants were then asked whether they believed that the cost of the machines used in the manufacturing process should influence their response, and whether or not the cost of the machines did influence their response.

Results

A total of 60 participants completed the study. Thirty participants completed the first version (high input for fudge A, low input for fudge B), and thirty participants completed the second version (low input for fudge A, high input for fudge B). On average, participants were 22.6 (5.71) years old, and just over half of the participants were male (53.7%).

Supporting our thesis, participants' evaluations of the fudge were significantly influenced by the expense of the machinery involved in making the fudge (see Fig. 3). When participants were told that the machinery used to

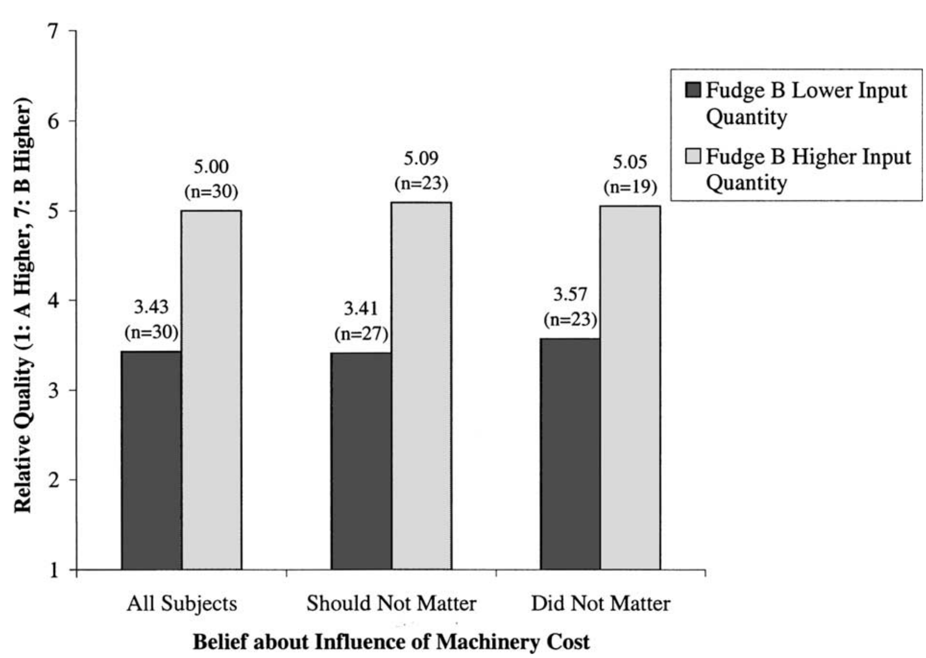

Fig. 3. Study 2—Fudge ratings by machine cost.

produce fudge A was more expensive than the machinery used to produce fudge B, they were significantly more likely to rate fudge A as higher quality, 3.43 (1.83) than when participants were told that the machinery used to produce fudge B was more expensive, 5.00 (1.39); $t(58) = 3.73$, $p < .001$. (Recall that low ratings indicate a belief that fudge A is higher quality than fudge B.)

Most participants (50 of 60, 83%) believed that the cost of the machinery should not influence their evaluations. Even when we consider evaluations of these 50 participants, we find the same pattern of results. When participants who believed that the machine cost should not matter were told that the machinery used to produce fudge A was more expensive than the machinery used to produce fudge B, they rated fudge A as higher quality than when they were told that the machinery used to produce fudge B was more expensive, 3.41 (1.85) versus 5.09 (1.41), $t(48) = 3.56$, $p < .001$.

Similarly, most participants (42 of 60, 70%) believed that the cost of the machinery did not influence their evaluations. When participants who believed that the machine cost did not matter were told that the machinery used to produce fudge A was more expensive than the machinery used to produce fudge B, they rated fudge A as higher quality than when they were told that the machinery used to produce fudge B was more expensive, 3.57 (2.06) versus 5.05 (1.54), $t(40) = 2.60$, $p = .01$.

Discussion

Results from this study extend our investigation of the input bias to a very different context. Participants in this study were asked to compare the relative quality of two samples of fudge. We consider a second type of input, the cost of machinery used to manufacture the fudge, and find that most participants believed this input was irrelevant to judging outcome quality. In spite of these beliefs, however, we again document the systematic misuse of input information in judgments of outcome quality.

Taken together, studies one and two identify an input bias. While input quantities are positively related to outcome quality in many cases, people may automatically associate high input quantities with high output quality—even when they recognize that input quantities should be irrelevant. We explore the mechanics of the input bias in study three.

Study 3

In study three we extend our investigation to explore the mechanics of the input bias. While we have assumed that decision makers automatically associate high input quantities with high outcome quality, in this study we consider an alternative mechanism—a selective encoding process. That is, participants in studies one and two may have viewed input information and formed expectations based upon this information. Subsequently, participants may have interpreted and encoded ambiguous outcome information in a manner consistent with their expectations (Jones, 1986).

To examine whether or not the input bias can be attributed to a selective encoding process, in study three we expose participants to input information after they experience the outcome. That is, unlike studies one and two in which participants were exposed to input information before experiencing the outcome, in study three participants learn about input information after experiencing the outcome. If the mechanism underlying the input bias is a selective encoding process, then input information in this study should not influence outcome assessments.

Methods

We use similar methods to those that we used in study one. In this study, however, participants only viewed one presentation (the electronic ink presentation), and instead of learning about the preparation time *before* viewing the presentation, they learned about the preparation time *after* viewing the presentation. We conducted two versions of this study. In the first version participants were told that the preparation time was long, 8 h 34 min (high input condition). In the second version participants were told that the preparation time was short, 37 min (low input condition).

As in study one we asked participants to evaluate the presentations along the same five dimensions: the overall quality of the presentation, the quality of the information presented, the presenter's knowledge of the subject, the organization of the presentation, and the quality of the presenter's presentation skills. Each evaluation was made on a 10-point response scale (1 = Poor, 10 = Excellent). As before, the reliability coefficients for both the long and short preparation time versions were very high ($\alpha = 0.91$ and $\alpha = 0.91$, respectively), and we use a composite measure of these five factors in the subsequent analysis.

Results

A total of 63 undergraduate participants completed the study (31 in the low input condition and 32 in the high input condition) as part of an optional class assignment. On average, participants were 21.4 (1.29) years old, and just over half of the participants were male (50.8%).

As in study one preparation time significantly influenced evaluations. Overall, ratings were significantly higher when preparation time was long than when preparation time was short, 7.65 (1.50) versus 6.21 (1.43), respectively; $t(61) = 3.90$, $p < .001$.

This same pattern of results characterizes ratings of participants who believed that preparation time should not influence their ratings ($n = 41$, 65%) and participants who believed preparation time did not influence their ratings ($n = 42$, 67%). As shown in Fig. 4, ratings for both groups were significantly higher when preparation time was long than when preparation time was short, 7.52 (1.49) versus 6.04 (1.35), $t(39) = 3.28$, $p < .01$, and 7.62 (1.37) versus 6.39 (1.42), $t(40) = 2.83$, $p < .01$. Only 24% of participants thought preparation time should matter and 24% of participants thought preparation time did matter (11% and 9.5%, respectively, were not sure).

Discussion

The methods we used in study three were similar to those we used in study one. In this study, however, participants learned about the preparation time after they viewed the presentation. Even with this change in methods, participants' evaluations were significantly higher when input quantities were high than when input quantities were low. This was true for participants who believed input quantity information should not matter and for participants who believed input quantity information did not matter. Taken together, results from this study suggest that the input bias is robust and subtle, and that a selective encoding process alone cannot account for this bias.

Study 4

In study four we extend our investigation to consider a boundary condition of the input bias. In this study we examine the moderating effect of outcome quality on the misuse of input information in judgments of outcomes.

Consistent with our results from study three and an association based mechanism (Arkes, 1991), we propose that people automatically associate high input quantities with high outcome quality. In this study we manipulate outcome quality in a way that fundamentally challenges this automatic association. Specifically, we examine the link between input quantity and perceived outcome quality for very low quality outcomes. We expect that when decision makers experience a very low quality outcome they will think more critically and less heuristically than they do when they experience a high quality outcome, and hence may disassociate high input quantities with good outcomes.

This proposition is related to an important finding in affect research. In this research scholars have found that people rely less on heuristics when they experience negative affect than when they experience positive affect (Bless, Bohner, Schwarz, & Strack, 1990; Bless et al., 1996; Isen, 1987; Mackie & Worth, 1989; Schwarz, 1990; Schwarz & Clore, 1983; Wegener, Petty, & Smith, 1995). Schwarz (1990) and Schwarz and Clore (1983) propose that negative affect states serve as signals that something in the decision maker's environment poses a problem. As a result, negative affect states motivate systematic information processing. Conversely, positive affect states do not heighten a decision maker's concern about

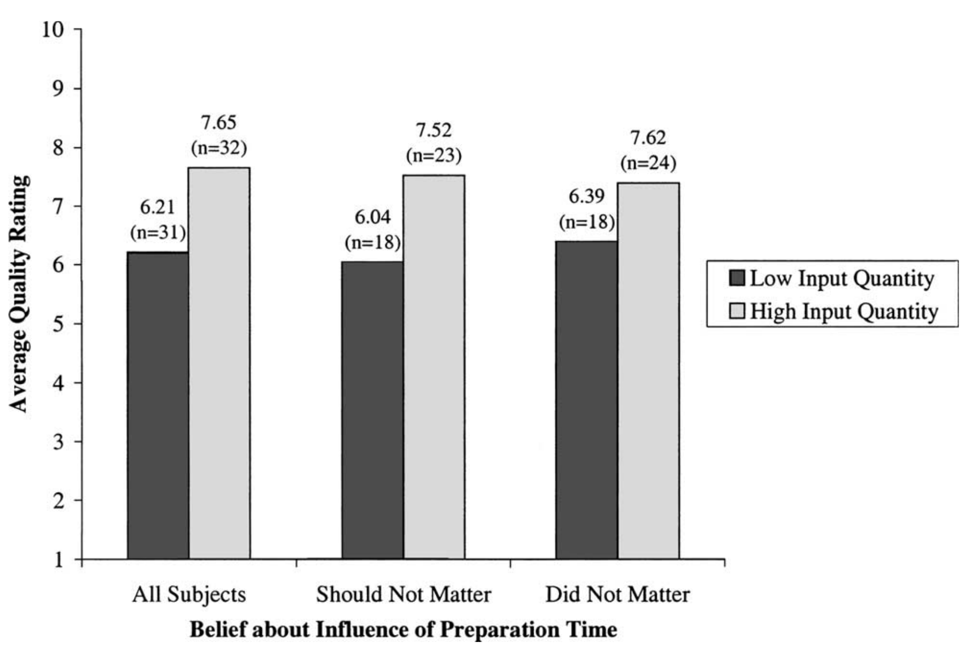

Fig. 4. Study 3—Electronic ink presentation ratings by input quantity: Input information presented after viewing presentation.

his or her environment, and as a result positive affect states lead to more heuristic information processing than negative affect states.

In this study we explore the operation of the input bias for both high and very low quality outcomes. We expect very low quality outcomes to motivate more systematic (and less heuristic) information processing than high quality outcomes, and as a result we expect the input bias to influence judgments of high quality outcomes more than judgments of very low quality outcomes. That is, in this study we consider the role of outcome quality in moderating the input bias, and we hypothesize a significant interaction between outcome quality and input information.

Methods

In study four we manipulated outcome quality. We conducted a taste test with ice tea, and asked participants to compare either two high quality samples or two low quality samples. Specifically, we asked participants to compare raspberry and lemon tea in one of two conditions—a high quality condition (where we added one third cup of sugar to two quarts of each kind of tea) and a low quality condition (where we added one tablespoon of salt and a half cup of lime juice to two quarts of each kind of tea).

We conducted a pilot study ($n = 9$) to validate our manipulation of high and low quality outcomes. For the pilot study the four samples were simply numbered one through four, and participants were asked to rate each of the four tea samples on a scale from 1 to 7 (1: very bad, 7: very good). A total of nine raters evaluated each of the four samples of ice tea. As expected, raters evaluated the teas with sugar (high quality) much higher than the teas with lime juice and salt (low quality). Average ratings for the high quality lemon tea, low quality lemon tea, high quality raspberry tea, and low quality raspberry tea were, 5.67 (0.5), 1.44 (0.53), 5.88 (0.78), and 1.67 (0.5), respectively. We conducted analysis of variance on these ratings, treating rater as a block, and the tea flavor and tea quality as fixed factors. The tea quality variable is highly significant. $F(1,8) = 420.07$, $p < .001$, indicating that the high quality teas were in fact much better than the low quality teas. No other variables in this analysis were significant ($p > .20$ for all other variables).

For the main study we recruited 160 participants to rate two samples of ice tea. These teas were labeled Tea A (which was always the lemon tea) and Tea B (which was always the raspberry tea). Participants were randomly assigned to one of four conditions from a 2 (Input) × 2 (Quality) design. The input conditions described either the lemon tea or the raspberry tea as having been made with an expensive machine. Specifically, participants were informed before their taste test that both samples of tea were "made with similar ingredients, however the ingredients were mixed and brewed differently. Tea A was made using [very expensive/inexpensive] machinery, and Tea B was made using [inexpensive/very expensive] machinery."

The quality condition was either high or low. In the high quality condition participants tasted high quality lemon tea (labeled Tea A) and high quality raspberry tea (labeled Tea B). In the low quality condition participants tasted low quality lemon tea (labeled Tea A) and low quality raspberry tea (labeled Tea B).

After tasting both teas, participants rated the two samples on a scale from 1 to 7 (1: Tea A is much higher quality, 7: Tea B is much higher quality). After responding to this question, participants were then asked whether they believed that the cost of the machines used in preparing the tea should influence their response, and whether or not the cost of the machines did influence their response.

Results

A total of 160 participants were recruited from a university campus food court. Participants were randomly and evenly assigned to each of the four conditions, and all 160 participants completed the study. On average, participants were 25.9 (10.05) years old, and just over half of the participants were male (54.8%).

Consistent with our prior results, preference ratings in the high quality conditions were higher when the raspberry tea had a high input description than when the raspberry tea had a low input description, 4.6 (1.74) versus 3.48 (1.61), respectively, $t(78) = 3.00$, $p < .01$. (Recall that high ratings indicate a belief that Tea B, the raspberry tea, is higher quality than Tea A, the lemon tea.) This pattern of results was true even among the 61 participants (76%) who believed that the cost of the machinery should not influence their ratings, 4.69 (1.71) versus 3.55 (1.43), $t(59) = 2.80$, $p < .01$, as well as among the 69 participants (86%) who believed that the cost did not influence their ratings, 4.67 (1.71) versus 3.53 (1.66), $t(67) = 2.80$, $p < .01$ (see Fig. 5).

This pattern of results, however, does not characterize ratings in the low quality conditions. In this case average ratings when the raspberry tea had a high input description were similar to the average ratings when the raspberry tea had a low input description, 4.28 (1.27) versus 4.68 (1.47), $t(78) = 1.30, p = n.s.$ This was also true among the 64 (80%) participants who believed that the machinery cost should not influence their ratings, 4.43 (1.30) versus 4.64 (1.50), $t(62) = .59$, $p = n.s.$, as well as the 60 participants (75%) who believed that the cost did not influence their ratings, 4.44 (1.20) versus 4.71 (1.45), $t(58) = .78$, $p = n.s.$ (see Fig. 6).

We conducted analysis of variance on participants' ratings as a function of outcome quality and input

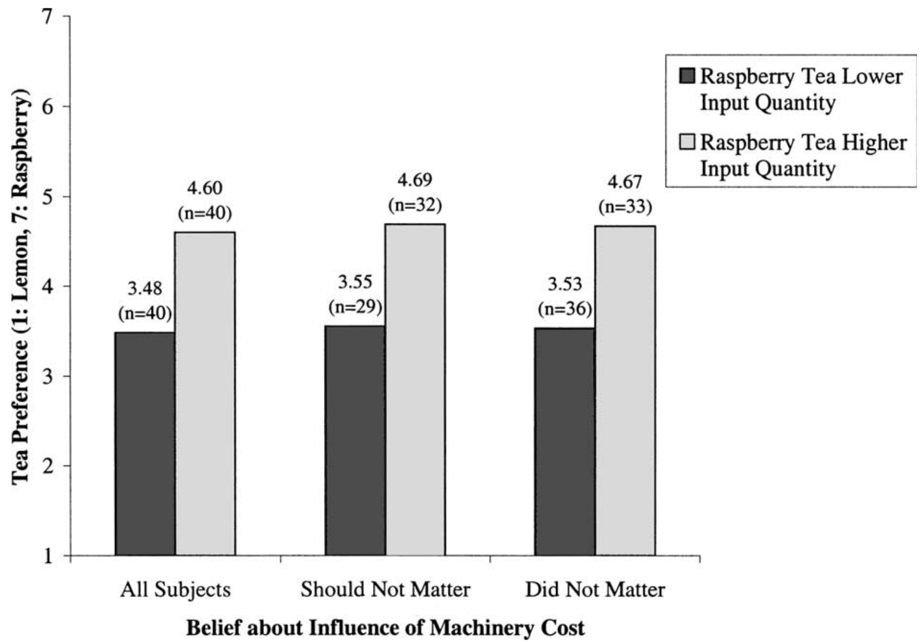

Fig. 5. Study 4—Preferences among high quality outcomes.

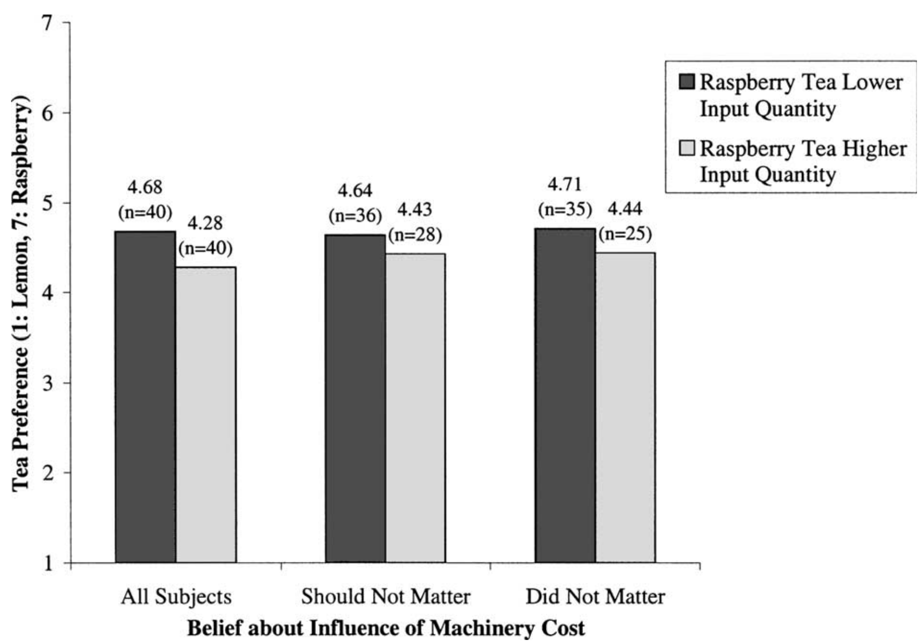

Fig. 6. Study 4—Preferences among low quality outcomes.

quantity. Supporting our hypothesis, we find a significant interaction between outcome quality and input quantity. This is true when we include all 160 participants, $F(1, 156) = 9.89$, $p < .01$, and when we include only those participants ($n = 125$) who believed that the cost of the machinery should not influence their ratings, $F(1, 121) = 6.244$, $p = .014$, as well as when we include only the 129 participants who believed that the cost of machinery did not influence their ratings, $F(1, 125) = 6.656$, $p = .011$.

Discussion

In this study we find a significant interaction between outcome quality and the input bias. While input information significantly biased evaluations of high quality outcomes, this same information did not influence evaluations of very low quality outcomes. Relative to high quality experiences, we believe low quality experiences heighten concern and motivate systematic information processing. In this case, decision makers who experience low quality outcomes are less likely to rely on heuristics, and as a consequence their judgments are less likely to be influenced by the input bias.

General discussion

Results from our studies identify a robust input bias. Irrelevant input information systematically influenced judgments of outcome quality. In our studies people evaluated outcomes more favorably when the inputs used to attain those outcomes were larger—even when they recognized that input information should not matter.

Consistent with Arkes' (1991) framework of judgment biases, we believe that people automatically associate high input quantities with high output quality. This association derives from the common observation that high input quantities are correlated with high output quality. Input quantity and output quality, however, are not always causally related, and participants in our studies appear to know this. Most of our respondents believed that input quantity (preparation time and productions costs) should *not* influence their evaluations.

Results from this work also examine the mechanism of the input bias. In study three we demonstrate that the input bias is not merely a function of selective encoding. Even when input information is presented *after* the evaluation process should have concluded, people are still influenced by input quantity information. In study four we extend our investigation to examine an important boundary condition of the input bias. In this study we find that irrelevant input information influenced judgments of high quality outcomes, but did not influence judgments of very low quality outcomes. We propose that relative to an experience with a high quality outcome, an experience with a very low quality outcome will cause decision makers to adopt a more systematic and less heuristic approach to evaluating outcomes.

Taken together, our findings have important implications for judgment and decision making. Outcome assessment is a key component of both individual and organizational decision making, and our results demonstrate that these assessments are likely to be systematically biased by the input quantities used—or purportedly used—to attain outcomes. In general, input measures are relatively easy to manipulate, and the misapplication of the input heuristic may be widespread.

Similar to Kerr's (1995) proposition that firms often reward for A (e.g., high production volume) while hoping for B (e.g., high production quality), we conjecture that managers often use input measures (e.g., the number of hours spent with a client) to assess productivity in a way that leads employees to make decisions that are not consistent with the firm's underlying goals.

The input bias is likely to be very difficult to correct. Across our studies we find a consistent input bias even among people who believe input information did *not* influence their judgment. According to both Arkes (1991) and Wilson and Brekke (1994) mere recognition of the bias will not correct this type of judgment error. In fact, even monetary incentives are unlikely to correct this bias (Arkes, 1991). Instead, for important decisions (e.g., promotion or hiring decisions) managers should use protocols that include blind review processes in which outcomes are judged without any knowledge of the inputs used to attain those outcomes.

A number of factors are likely to exacerbate the input bias. In general, we expect the input bias to exert the most influence when objective outcome values are difficult to evaluate and objective input criteria are salient and easy to measure (Eisenhardt, 1985; Ouchi, 1979). This proposition is related to prior work that has found that subjective evaluations of ambiguous outcomes are often influenced by irrelevant factors (Hsee, 1995; Hsee, 1996; Marshall & Mowen, 1993; Schweitzer & Hsee, 2002). When an outcome lacks objective criteria other factors may influence evaluations. For example, Hsee (1996) demonstrates that self-interest influences judgment more when outcomes are ambiguous.

A second factor likely to moderate the input bias is accountability. For example, asking people to justify their judgments may actually increase their reliance on input measures as people search for objective (though potentially irrelevant) measures to support their evaluations. This proposition is related to Tetlock and Boettger's (1989) finding that accountability magnifies the dilution effect.

The input bias is likely to have broad implications for managerial decision making. In many domains input quantities are easy to manipulate or misrepresent. At the same time, the input bias may be particularly difficult to debias. Our results identify a robust link between input measures and judgments of outcomes—even when people recognize that input information should not matter and believe that input information did not matter. Future work should investigate the effectiveness of debiasing techniques, and identify implications of this bias for the structure of incentive systems, organizational design, and management practice.

Acknowledgments

We thank Jon Baron, Gerard Cachon, Terry Connolly, Jack Hershey, Howard Kunreuther, and three anonymous reviewers for their helpful comments and suggestions.

References

Arkes, H. R. (1991). Costs and benefits of judgment errors: Implications for debiasing. *Psychological Bulletin, 110*, 486–498.

Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes, 35*, 124–140.

Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology, 54*, 569–579.

Baron, J., & Maxwell, N. P. (1996). Cost of public goods affects willingness to pay for them. *Journal of Behavioral Decision Making, 9*, 173–183.

Bastardi, A., & Shafir, E. (1998). On the pursuit and misuse of useless information. *Journal of Personality and Social Psychology, 75*, 19–32.

Baysinger, B. D., Kosnik, R. D., & Turk, T. A. (1991). Effects of board and ownership structure on corporate strategy. *Academy of Management Journal, 34*, 205–214.

Bazerman, M., Giuliano, T., & Appelman, A. (1984). Escalation of commitment in individual and group decision making. *Organizational Behavior and Human Performance, 33*, 141–152.

Bless, H., Bohner, G., Schwarz, N., & Strack, F. (1990). Mood and persuasion: A cognitive response analysis. *Personality and Social Psychology Bulletin, 16*, 331–345.

Bless, H., Clore, G. L., Schwarz, N., Golisano, V., Rabe, C., & Wolk, M. (1996). Mood and the use of scripts: Does a happy mood really lead to mindlessness? *Journal of Personality and Social Psychology, 71*, 665–679.

Camerer, C., Loewenstein, G., & Weber, M. (1989). The curse of knowledge in economic settings: An experimental analysis. *Journal of Political Economy, 97*, 1232–1254.

Eisenhardt, K. (1985). Control: Organizational and economic approaches. *Management Science, 31*, 134–149.

Fischer, G., Damodaran, N., Laskey, K., & Lincoln, D. (1987). Preferences for proxy attributes. *Management Science, 33*, 198–215.

Garland, H. (1990). Throwing good money after bad: The effect of sunk costs on the decision to escalate commitment to an ongoing project. *Journal of Applied Psychology, 75*, 728–731.

Graves, S. B. (1988). Institutional ownership and corporate R & D in the computer industry. *Academy of Management Journal, 31*, 417–428.

Hansen, G. S., & Hill, C. W. L. (1991). Axe institutional investors myopic. A time-series study of four technology-driven industries. *Strategic Management Journal, 12*, 1–16.

Hawkins, S., & Hastie, R. (1990). Hindsight biased judgments of past events after the outcomes are known. *Psychological Bulletin, 107*, 311–327.

Hsee, C. (1995). Elastic justification in decision making: How task irrelevant but tempting considerations influence decisions. *Organizational Behavior and Human Decision Processes, 62*, 330–337.

Hsee, C. (1996). Elastic justification: How unjustifiable factors influence judgments. *Organizational Behavior and Human Decision Processes, 66*, 122–129.

Hsee, C., Blount, S., Loewenstein, G., & Bazerman, M. (1999). Preference reversals between joint and separate evaluations of options: A review and theoretical analysis. *Psychological Bulletin, 125*, 576–590.

Isen, A. M. (1987). Positive affect, cognitive processes, and social behavior. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 20, pp. 203–253). New York: Academic Press.

Jones, E. E. (1986). Interpreting interpersonal behavior: The effects of expectancies. *Science, 234*, 41–46.

Josephs, R. A., Giesler, R. B., & Silvera, D. H. (1994). Judgment by quantity. *Journal of Experimental Psychology: General, 123*, 21–32.

Kerr, S. (1995). On the folly of rewarding A, while hoping for B; More on the folly. *Academy of Management Executive, 9*(1), 7–16.

Levin, R., Cohen, W., & Mowery, D. (1985). R & D appropriability, opportunity and market structure: New evidence on some Schumpeterian hypotheses. *American Economic Review, 75*(2), 20–24.

Mackie, D. M., & Worth, L. T. (1989). Processing deficits and the mediation of positive affect in persuasion. *Journal of Personality and Social Psychology, 57*, 27–40.

March, J. G. (1987). Ambiguity and accounting: The elusive link between information and decision making. *Accounting, Organizations and Society, 12*, 153–169.

Marshall, G., & Mowen, J. (1993). An experimental investigation of the outcome bias in salesperson performance evaluations. *Journal of Personal Selling and Sales Management, 13*(3), 31–47.

Moffat, S. (1990). Japan's new personalized production. *Fortune*, 132–135.

Nisbett, R. E., Zukier, H., & Lemley, R. E. (1981). The dilution effect: Nondiagnostic information weakens the implications of diagnostic information. *Cognitive Psychology, 13*, 248–277.

Ouchi, W. (1979). A conceptual framework for design of organizational control mechanisms. *Management Science, 25*, 833–848.

Schoenecker, T. S., Daellenbach, U. S., & McCarthy, A. M. (1995). Factors affecting a firm's commitment to innovation. *Academy of Management Best Papers Proceedings, 1995*, 52–56.

Schwarz, N. (1990). Feelings as information: Informational and motivational functions of affective states. In R. M. Sorrentino, & E. T. Higgins (Eds.), *Handbook of Motivation and Cognition: Foundations of Social Behavior* (Vol. 2, pp. 527–561). New York: Guilford.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology, 45*, 513–523.

Schweitzer, M., & Hsee, C. (2002). Stretching the truth: The role of self-justification in the deception decision process. *Journal of Risk and Uncertainty, 25*, 185–201.

Scott, L. (1997). Fighting to keep a secret. *Modern Healthcare, 27*, 30.

Silvera, D. H., Josephs, R. A., & Giesler, R. B. (2001). The proportion heuristic: Problem set size as a basis for performance judgments. *Journal of Behavioral Decision Making, 14*, 207–221.

Staw, B. (1976). Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action. *Organizational Behavior and Human Performance, 16*, 27–44.

Tetlock, P. E., & Boettger, R. (1989). Accountability: A social magnifier of the dilution effect. *Journal of Personality and Social Psychology, 57*, 388–398.

Thaler, R. (1985). Mental accounting and consumer choice. *Marketing Science, 4*(3), 199–214.

Wegener, D. T., Petty, R. E., & Smith, S. M. (1995). Positive mood can increase or decrease message scrutiny: The hedonic contingency view of mood and message processing. *Journal of Personality and Social Psychology, 69*, 5–15.

Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin, 116*, 117–142.

Received 21 March 2002