Do Financial Statement Users Judge Relevance Based on Properties of Reliability?

Lisa Koonce

McCombs School of Business
University of Texas at Austin

Date: Friday 11th March
Time: 3:00 to 4:30 pm
Venue: Room 2063 Quad.
Do Financial Statement Users Judge Relevance Based on Properties of Reliability?

Kathryn Kadous  
Emory University

Lisa Koonce  
The University of Texas at Austin

Jane Thayer  
University of Georgia

October 2010

Abstract: Relevance and reliability (now referred to as “representational faithfulness”) are qualities of financial information that both the Financial Accounting Standards Board and the International Accounting Standards Board use in setting standards for financial reporting. Despite their importance, very little research has addressed how financial statement users use these constructs. Although standard setters view relevance and reliability as independent, we show via three experiments that users do not view these two constructs as independent. Instead, users’ assessments of the relevance of an economic construct are influenced by variations in properties that are associated with the reliability of its measurement. The relationship between assessed relevance and reliability is unidirectional, in that factors underlying reliability influence judgments of relevance, but factors underlying relevance do not influence judgments of reliability. Our findings are important because inappropriate assessments of relevance can influence firm valuation. The results are particularly meaningful in the context of fair value because such measurements can vary widely in reliability.

JEL Classification: M41  
Keywords: Relevance, Reliability, Fair Value, Valuation.

The authors appreciate the helpful comments of brownbag participants at Emory University and The University of Texas at Austin, as well as workshop participants at the University of Florida, University of Georgia, University of Oklahoma, and Florida State University. We appreciate the helpful comments of Tina Carpenter, Margaret Christ, Jeffrey Hales, Jackie Hammersley, Frank Hodge, Susan Jurney, Molly Mercer, and James Smith. We also appreciate the financial
assistance of the Deloitte Foundation. This paper has also benefitted from conversations with Ross Jennings and Katherine Schipper.

1. Introduction

Relevance and reliability (now referred to as "representational faithfulness") are important qualities of financial reporting that both the Financial Accounting Standards Board (FASB) and the International Accounting Standards Board (IASB) use in setting standards (FASB 2010; IASB 2010).¹ Relevance addresses the pertinence of an economic construct (e.g., fair value, historical cost) to a user’s decision. Reliability addresses how well that economic construct, or phenomenon, is depicted or measured (e.g., fair value based on a market transaction or model inputs). Standard setters state that although the decision usefulness of accounting information is a joint function of its relevance and reliability, these two qualities are distinct and independent (FASB 2010). That is, relevance should be judged without consideration of the manner in which the economic construct is depicted (Maines and Wahlen 2006; FASB 2010). However, anecdotal evidence suggests that financial statement users may not think about these constructs in this fashion (Schipper 2003, 66).

Investigating how financial statement users think about the constructs of relevance and reliability is important for several reasons. First, we lack empirical evidence regarding how users assess these qualities. Even the limited research and debates to date suggest that the answer is unclear. Some researchers indicate that users view relevance and reliability as inversely related (e.g., Schipper 2003), while others assert a positive relationship (McCaslin and Stanga 1993). Second, although the purpose of the conceptual frameworks is to guide standard

¹ The FASB and IASB recently replaced the term reliability with representational faithfulness because they were concerned that the auditing profession associated reliability both with the factors it was intended to capture and also with a factor it was not intended to capture—vouchability (Schipper 2005). This modification to the conceptual framework does not impact our study because we examine features of reliability (i.e., error and bias) that are unrelated to the change. We use the more familiar term reliability herein.
setting, users’ understanding of the characteristics defined in the frameworks is nevertheless important. That is, acceptance of FASB and IASB standards is fundamentally tied to users’ appreciation for the basis of those standards. For example, some of the criticisms raised against fair value accounting during the recent financial crisis can be tied, in part, to users confusing the relevance of fair values with their reliability (American Bankers Association 2009).

In this paper, we report the results of three experiments in which we investigate the hypothesis that financial statement users’ assessments of the relevance of an economic construct will be influenced by factors (i.e., potential for bias and error) underlying the reliability of its measurement.2 We examine this idea in the context of fair value measurement of an asset, as relevance and reliability issues are of particular practical concern there.3 We also test the proposition that this judged association is unidirectional. That is, we posit that manipulations of factors underlying the reliability of a measurement will influence assessments of the reported construct’s relevance (and, of course, measurement reliability). However, we do not expect the converse to be true. That is, we do not expect that manipulations of factors underlying relevance will affect users’ assessments of reliability. Finally, we also investigate the economic importance of our ideas by determining whether users’ judgments about a firm’s valuation are influenced by the relationships noted above.

Our hypotheses are grounded in psychology theories about attribute substitution and representativeness. These theories predict that when individuals are asked a question they do not know how to answer, they often substitute an easier-to-answer question in its place (Kahneman and Frederick 2002). We propose that when asked the difficult question “How relevant is this

---

2 The purpose of this study is not to test users’ understanding of the FASB and IASB definitions of relevance and reliability per se, but to examine how users interpret and use the constructs that these terms represent.

3 Although our paper is set in the context of fair value, it does not address the issue of whether fair value is more or less relevant than historical cost (see Koonce et al. 2010 for a study of that issue). Rather, our study focuses on how the judged relevance of fair value is influenced by variations in the reliability of its measurement.
item?” financial statement users will, instead, answer an easier question such as “How good is the item’s measurement?” This particular substitution appears plausible as prior research has established that reliability is easier to assess (and is more accurately assessed) than relevance (Jurney 2008; Koonce, et al. 2010). As a result, users are likely to judge the relevance of a reported economic construct by appraising how well it represents, or is similar to, a good measurement (i.e., by use of the representativeness heuristic, Tversky and Kahneman 1974).

The results of our experiments support our hypothesis that variations in factors associated with the reliability of an asset’s fair value measurement affect financial statement users’ judgments of the relevance of fair value. Interestingly, we observe that not all factors underlying reliability have this effect. We find that users’ judgments about relevance are strongly influenced by the competence of the source of the measurement and by whether the measurement is based on a market transaction versus a model. However, relevance judgments are not strongly affected by our manipulation of bias in the information source.

We also observe evidence for the hypothesized unidirectional nature of the relationship between assessments of relevance and reliability. While properties associated with reliability affect users’ assessments of relevance, properties associated with relevance do not influence users’ assessments of reliability. Furthermore, results support the idea that factors underlying reliability appear to be double counted in users’ valuation judgments—that is, these properties influence valuation through assessments of both reliability and relevance. This result suggests that properties underlying reliability may be over-weighted in valuing a firm.

Our results have several important implications. First, our results imply that efforts of the FASB and the IASB to extend the use of fair values in financial reports are likely to be met with resistance. We show that acceptance of the concept of fair value measurement of assets and
liabilities may be conditioned primarily on how well the items are measured rather than on separate assessments of the relevance of the economic construct of fair value and the reliability of its measurement. This idea finds anecdotal support with the recent financial crisis, during which observers argued that thin markets and unreliable measurements created fair value accounting outcomes that were not relevant to decision makers (American Bankers Association, 2009). We believe our study is important to this debate as it indicates that at least part of the reason users argue against the construct of fair value may be that they confuse or conflate reliability and relevance.

Second, users’ confusion about the qualities of reliability and relevance may hamper additional attempts to improve the actual reliability of fair value measures or the communication of reliability. Although the FASB has previously issued guidance about the communication of fair value measurement reliability (e.g., SFAS 157), further efforts to improve those disclosures may be thwarted if users are unable to accept the relevance of the fair value construct independent of the reliability of its measurement. By documenting that users do not view separately these two qualities, we hope that our study provides a foundation for greater clarity on what can be improved (i.e., measurement reliability or its disclosure) and what is relatively fixed (i.e., relevance).

Third, our study provides initial insight into which properties underlying reliability create confusion for financial statement users. Specifically, the expectation of bias in the fair value measurement did not strongly influence perceptions of the relevance of fair value information. However, the expectation of error in the measurement and the use of a model versus a market measure (which, per Song et al. (2010), may represent a joint manipulation of error and bias) had strong effects. This finding may suggest that recent concerns about bias in the computation of
fair value measurements undermining their decision usefulness are misplaced (Zion, et al. 2008,
2009). 4

Fourth, our results suggest that financial statement users’ conflation of relevance and
reliability may influence their valuation judgments (i.e., strengths or deficits in reliability due to
potential error appear to be double-counted in valuations). The natural next question is whether
this confusion can be remedied. As this confusion appears to arise from the use of a heuristic,
training in assessing relevance may not be sufficient to correct decision errors (e.g., see Tversky
and Kahneman 1974). Correcting these errors may require providing users with a decision aid in
the form of a framework or process for making and using assessments of relevance.

In the following section, we develop the theory that leads to our theoretical predictions.
Sections three, four, and five describe our three experiments, respectively. Section six
summarizes and concludes the paper.

2. Hypotheses Development

Relevance and reliability are the two primary qualitative characteristics of accounting
information that determine its usefulness in decision making (FASB 1980, 2010). Relevance is
concerned with the pertinence of an economic construct to financial statement users’ decision
making. Reliability, on the other hand, is concerned with the depiction, or measurement, of the
economic construct. Both relevance and reliability influence the reported measure’s decision
usefulness, but the level of reliability of a measure does not influence the relevance of an
economic phenomenon.

4 The archival literature is mixed regarding whether bias in fair values is problematic for users. Barth and Clinch
(1998) find no difference in the value relevance of independent appraiser-based asset revaluation estimates and
management-based estimates among Australian companies. In contrast, Muller and Riedl (2002), in their study of
UK investment property firms, find that market participants judge independent appraiser-based estimates as more
reliable.
An example outside of the financial reporting context helps illustrate these ideas. Consider an individual who is deciding whether to wear a coat. The relevant phenomenon to the individual’s decision is the outside temperature. A measure, or depiction, of the outside temperature can be gleaned from a thermometer reading. If the thermometer reading is unreliable, it is not useful to the individual’s decision of whether to wear a coat. However, the unreliable measure does not negate the relevance of the outside temperature to the decision of whether to wear a coat. As shown by this example, although a lack of either relevance or reliability precludes a measure from being decision useful, the qualities of relevance and reliability are independently determined.\(^5\)

While the FASB and IASB indicate that relevance and reliability are distinct, we question whether financial statement users treat them as such. As explained more fully below, prior research indicates that financial statement users have fairly accurate ideas about how to assess reliability, and so we expect they will be able to do so readily. On the other hand, users are unlikely to have a ready framework for assessing relevance, and so they will rely on the representativeness heuristic (Tversky and Kahneman 1974), by which factors underlying reliability will intrude on assessments of relevance (Joyce et al. 1982; Jurney 2008; Koonce et al. 2010).

Reliability is a basic property of information, and we expect that individuals are well acquainted with its determinants. To support this idea, we note that numerous studies in accounting and psychology provide evidence that individuals understand that factors underlying reliability, such as the objectivity of the measurement source and the error-proneness of that source, impact the decision usefulness of a reported measure. For example, Hirst (1994) showed that auditors discount the inferential value of information obtained from sources that are low in

\(^5\) We thank Molly Mercer for this example.
competence or objectivity relative to that of information obtained from sources higher in these characteristics. Similarly, market participants discount forecasts made by management with low prior accuracy in forecasting earnings relative to those with high prior accuracy (Williams 1996; Hirst et al. 1999). Moreover, market participants’ reactions to management forecasts are, in part, a function of the credibility of management (Jennings 1987). Investors glean manager credibility from factors such as their competence, trustworthiness, and situational incentives (Mercer 2004). Additional studies indicate that users of other forms of financial reporting can identify when a source is more or less credible and will utilize the information accordingly. For example, users base assessments of financial analysts’ credibility in part on analysts’ prior accuracy, and users’ willingness to purchase future reports from these analysts is associated with their assessments of credibility (Kadous et al. 2009).

Unlike reliability, the concept of relevance is situation-specific. The relevance of a measurement depends on the specific decision to be made and its context (FASB 1980, 2010). Because of this uniqueness, users are less likely to have an accurate framework or schema for assessing relevance. Consistent with this thinking, recent studies in the context of fair value measurements confirm that financial statement users assess reliability with greater ease than they do relevance. Both Jurney (2008) and Koonce et al. (2010) find that investors correctly identify characteristics that are typically associated with reliability, but they have a more difficult time understanding the concept of relevance (also see Joyce et al. 1982). In these studies, participants appear to conflate properties underlying reliability (e.g., representational faithfulness, measurement error) with the quality of relevance.6

---

6 While Jurney (2008) and Koonce et al. (2010) demonstrate users’ relative difficulty in assessing relevance versus reliability, neither addresses why this occurs and neither tests our hypotheses that variations in properties associated with reliability influence relevance judgments but that the converse is not true (i.e., variations in relevance do not affect reliability judgments).
Because users are unlikely to have a readily available framework for assessing relevance, we expect they will make the judgment intuitively, using heuristics (e.g., Tversky and Kahneman 1974). A core idea underlying decision heuristics is that of attribute substitution—when asked a difficult question, individuals answer an easy one instead (Kahneman and Frederick 2002). We propose that when asked the difficult question “How relevant is this item?” financial statement users will, instead, answer an easier question such as “How good is the item’s measurement?” Thus, they will assess the relevance of a reported economic construct by assessing how well it represents, or is similar to, a good measurement (i.e., by use of the representativeness heuristic, Tversky and Kahneman 1974).

To the extent that financial statement users view reliability, a relatively easy-to-assess characteristic, as the standard for a good measurement, users will view a relevant reported economic construct as having been generated by a reliable measurement process or source. This reasoning implies that financial statement users will correctly use factors underlying the reliability of a measurement in assessing its reliability, but they will also incorrectly rely on those same variables in assessing relevance. This logic leads to our first hypothesis.

H1: Financial statement users’ assessments of the relevance of fair value information will be influenced by factors underlying the reliability of its measurement (i.e., potential for bias and error).

Although our theory predicts that assessments of reliability and relevance will be positively associated, it is important to note that we also expect that the relationship is unidirectional. Manipulations of factors underlying the reliability of a measurement will influence assessments of the reported construct’s relevance (and, of course, measurement reliability). However, we do not expect the converse to hold—that is, manipulations of factors underlying relevance will not
affect users’ assessments of reliability. The latter is unlikely to occur because users are more conversant with assessing reliability and will not rely on heuristics to do so.

The decision usefulness of information is a function of both its relevance and its reliability (FASB 1980, 2010). Thus, we expect users will rely on their own assessments of the relevance and reliability of information in determining how useful that information is for subsequent decisions, such as valuation. If variations in properties underlying reliability affect assessments of reliability, as is appropriate, but also affect assessments of relevance as we predict in our first hypothesis, and if users incorporate assessments of both relevance and reliability into their valuation judgments, then variations in properties underlying reliability will be double-counted in valuing a firm. In other words, variations in properties underlying the reliability of a measurement will influence valuation through two paths—through users’ assessments of the relevance of the item and through their assessments of the reliability of the item’s measurement. This prediction is summarized in the following hypothesis.

H2: Financial statement users’ valuation judgments will be influenced by factors underlying the reliability of a measurement via two effects—one through assessed reliability and a second through assessed relevance.

It is difficult to establish a normative benchmark for weights that users should apply to various factors when making valuation judgments. However, a reasonable interpretation of a finding that properties underlying reliability influence valuation via two paths—one appropriate and one not—is that those properties are likely over-weighted in valuation.
3. **Experiment 1**

*Participants, Design, and Variables*

We tested our hypotheses in an experiment with 129 MBA students from two *Business Week* top-25 schools. Participants were asked to assume the role of a potential investor in a manufacturing firm. All were told that the company had purchased a plot of land for $41 million two years ago. The current value of the land, generated by external consultants who relied on a model, was estimated at $53 million. As explained more fully below, we manipulated two variables in a $2 \times 2$ full-factorial, between-participants design, relating to properties underlying the reliability and relevance of the land’s current value. Participants were randomly assigned to a condition.

We manipulated the competence of the source (i.e., external consultants) and held constant the source’s bias at a low level. We varied competence at two levels. Those participants in the high competence (i.e., high reliability) conditions were provided with the following information regarding the consultants:

> Based on your experience and knowledge, you believe the consultants are unbiased, and based on the accuracy of their work you have noted on other valuations, you believe they are highly competent.

Participants in the low competence (i.e., low reliability) condition were provided with the following statement:

---

7 Participants came from three sources: 24 were full-time MBA students from one institution (average of 4.7 years of work experience, 4.9 finance classes, 3.1 accounting classes, 71% had invested in individual stocks previously and 96% plan to do so), 55 were full-time MBA students from a second institution (average of 5.5 years work experience, 2.5 finance classes, 2.1 accounting classes, 65% had invested in stocks previously and 91% plan to), and 50 were executive MBA students from the first institution. Time constraints prevented us from collecting individual demographics from the executive students; however, the class as a whole averaged 12 years of work experience. Source of participants did not have significant effects on any of our dependent measures.

8 This manipulation represents higher and lower levels of reliability rather than absolute levels of high and low. For ease of exposition, though, we refer to levels of this variable, as well as other analogous variables in the paper, as “high” and “low” rather than “higher” and “lower.”
Based on your experience and knowledge, you believe the consultants are unbiased, but, based on errors and reasoning flaws you have noted on other valuations, you believe they are not very competent.

Our second manipulation relates to a property we expect financial statement users to associate with relevance—the length of time until management plans to sell the asset (FASB 2006, SFAS 157, ¶20). Although there is debate as to whether this property should affect relevance judgments in a valuation context (Hague and Willis 1999), it appears to be a common factor influencing such judgments (Oster 2003; Greenberg 2008). We manipulated the time to sale of the asset at two levels. Approximately half of the participants were told that management intends to sell the plot of land very soon (“sell soon” conditions), which corresponds to higher relevance for the current valuation decision we assigned participants. The remaining participants were told that management does not intend to sell the plot of land for at least 20 more years (“no sell” conditions), which corresponds to lower relevance for this decision. Figure 1 provides an example of the materials.

[Insert Figure 1]

After being presented with the information described above, participants responded to five questions. First, they were asked to assess the relevance of the current value information to their evaluation of the firm. Participants responded using a 101-point scale, with 0 labeled “not at all relevant” and 100 labeled “extremely relevant.” The responses to this question serve as the main dependent measure for our first hypothesis. The second question determines whether these relevance judgments translate into differences in judged firm value, our second dependent measure. Specifically, participants assessed whether the situation described in the case materials increases, decreases, or neither increases nor decreases the value they would place on the firm. Participants who indicated a change in the value were asked to indicate the magnitude of the
change on a 101-point scale, ranging from “slightly increases (decreases) how much I value the company” (0) to “greatly increases (decreases) how much I value the company” (100). We created our valuation dependent variable by coding the magnitude of change as a positive amount for increases in value and a negative amount for decreases in value. This variable is zero for participants who indicated neither an increase nor a decrease in valuation due to the current value information.

The third question is a measure to determine whether our manipulation of competence influenced participants’ judgments of the measurement’s reliability. The question asks participants to assess how reliable the current value number is on a 101-point response scale ranging from 0, “very unreliable,” to 100, “very reliable.” Questions four and five are control measures of participants’ assessments of the amount of bias in the current value measurement and management’s forthcomingness.

Results

Preliminary Tests. We first assess whether our manipulation of competence influenced participants’ judgments of the reliability of the current value number. Participants in the high competence conditions assessed the reliability of the current value number significantly higher (mean of 67.14) than did those in the low competence conditions (mean of 45.39) ($F_{1,125} = 42.05$, $p < 0.01$). As expected, neither time to sale nor the interaction of time to sale and competence significantly influenced reliability (both $p$-values > 0.30). Also as expected, participants’ ratings of bias and management forthcomingness did not significantly differ across levels of competence (both $p$-values > 0.10).

---

9 An ANOVA model for bias revealed a marginally significant main effect of time to sale ($p = 0.085$) and a significant interaction of competence and time to sale ($p < 0.01$). While bias was rated as approximately equal in the two higher competence conditions, bias was rated as higher for lower competence sources when management had no plans to sell and lower for lower competence sources when management planned to sell the asset soon.
Tests of Hypotheses. Descriptive (inferential) statistics for participants’ assessments of relevance and their valuation judgments are presented in Panel A (Panel B) of Table 1. Turning first to the relevance dependent measure, the analysis of variance (ANOVA) model reveals a significant main effect of the competence variable \((F_{1,125} = 10.92; p < 0.01)\), but no other significant effects. Cell means indicate that participants receiving the measurement from a high competence source viewed the current value information as significantly more relevant (mean of 65.66) than did participants receiving the measurement from a low competence source (mean of 51.44). Thus, financial statement users assess the current value measurement with higher reliability as more relevant, supporting Hypothesis 1.

As noted above, there were no significant effects in the model for participants’ relevance ratings other than that of source competence. The manipulation of time to sale did not have the expected effect on ratings of relevance. Thus, we have evidence that competence influences users’ assessments of both reliability and relevance and that time to sale does not impact reliability. However, because our manipulation of time to sale did not also affect assessed relevance, we cannot be sure that we have effectively manipulated factors underlying relevance. Therefore, Experiment 1 allows us to test for manipulations of reliability influencing relevance judgments, but not the converse.

[Insert Table 1]

Participants’ valuation judgments are also provided in Table 1. The ANOVA results show a significant main effect of competence on valuation effects \((F_{1,114} = 8.46, p < 0.01)\). Consistent with the relevance assessments, participants receiving the measurement from a high competence source indicated more positive valuation effects (mean of 34.24) than did those receiving the
measurement from a low competence source (mean of 16.12). Time to sell and the interaction of competence and time to sell are not significant in the ANOVA.

Hypothesis 2 examines the paths by which properties underlying reliability influence valuation judgments. This hypothesis is important for two reasons. First, although we find support for Hypothesis 1, it is possible that the relationship between manipulated competence and assessed relevance is not meaningful in the sense of influencing valuation. Second, Hypothesis 2 verifies that users incorporate both assessed reliability and relevance in their valuations, and that properties underlying reliability are double-counted in valuation judgments. We test our theory underlying both Hypothesis 1 and 2 using a structural-equations model, as presented in Figure 2.

[Insert Figure 2]

Overall, the model fits the data well.\(^\text{10}\) The traditional $\chi^2$ test shows a good fit ($\chi^2 = 0.99, p = 0.32$), as do other standard fit measures.\(^\text{11}\) Turning to the path coefficients, the relationship between our manipulated competence variable and perceived reliability (Link 1) is positive and significant ($p < 0.01$), consistent with a successful manipulation of factors underlying reliability. The relationship between competence of the valuation source and perceived relevance (Link 2) is positive and significant ($p < 0.01$), consistent with Hypothesis 1. Finally, both assessed reliability (Link 3) and assessed relevance (Link 4) are, in turn, positively associated with the valuation effect (both $p$-values < 0.05). These results support our theory that variations in competence cause users to assess differences in reliability, as one would expect, but also in

\(^{10}\) Our full hypothesized model would also include a link from manipulated factors influencing relevance (time to sale) to assessed relevance. However, because the manipulation of time to sale did not influence participants’ assessments of relevance, we omit this variable from the model.

\(^{11}\) The Comparative Fit Index, a measure of the proportion of improvement of the fit of our model to the null model, is 1.00, which is above the generally accepted minimum value of 0.95 (Byrne 2001). The Incremental Fit Index (1.00) is above the recommended minimum of 0.95 (Byrne 2001), and the Root Mean Square Error of Approximation (0.00) is below the recommended maximum of 0.01.
relevance. In addition, both assessed reliability and relevance then affect valuation judgments.

Both of the indirect effects of competence on valuation (through judged reliability and through judged relevance) are significant ($p < 0.01$) by the product of Z-Scores test (Holbert and Stephenson 2003). These results support Hypothesis 2.

4. Experiment 2

Participants, Design, and Variables

Experiment 2 was conducted at the same time as Experiment 1 and is identical to it with one exception. Instead of varying the competence of the source of the current value measurement, we manipulated the expected bias in that source. Accordingly, in Experiment 2, we held source competence constant at a high level and we manipulated bias at two levels, high or low. As in Experiment 1, we manipulated the time to sale of the asset (sell soon versus no sell) as a factor underlying relevance. Accordingly, our design is a $2 \times 2$ full-factorial, between-participants design. Participants in Experiment 2 are 131 MBA students from two Business Week top-25 schools.

Approximately half of the participants were randomly assigned to a condition in which they were told that firm management had derived the current value of the asset (i.e., high bias). These participants read the following statement:

*Based on your experience and knowledge, you believe management is highly competent, but based on errors in its prior estimates, you believe management has biased views of its assets’ values.*

---

12 Experiments 1 and 2 were conducted at the same time using the same populations of participants. Because of the similarity of the two experiments, we were able to use the two high competence cells from Experiment 1 in our second experiment. Those two cells are the low bias cells in Experiment 2.

13 Experiment 2 participants came from three sources: 26 were full-time MBA students from one institution (average of 5.8 years of work experience, 4.0 finance classes, 2.6 accounting classes, 69% had invested in individual stocks previously and 92% plan to do so), 55 were full-time MBA students from a second institution (average of 5.5 years work experience, 2.5 finance classes, 2.0 accounting classes, 64% had invested in stocks previously and 89% plan to), and 50 were executive MBA students from the first institution (this class as a whole averaged 12 years of work experience). Source of participants did not have significant effects on any of our dependent measures.
The remaining participants were assigned to a condition in which they were told that the current value had been determined by external consultants (i.e., low bias). They read the following:

*Based on your experience and knowledge, you believe the consultants are unbiased, and based on the accuracy of their work you have noted on other valuations, you believe they are highly competent.*

As previously noted, time to sale was manipulated as in Experiment 1. Response measures were the same five items captured in Experiment 1.

**Results**

**Preliminary Tests.** We first assess whether our manipulation of bias influenced participants’ reliability judgments. Participants in the low bias conditions rated the reliability of the measurement higher (mean of 67.14) than did participants who were told the measurement was prepared by the high bias source (mean of 50.50) \((F_{1,127} = 27.38, p < 0.01)\). No other significant effects were observed, as expected \((p\text{-values} > 0.30)\). We also observed a significant main effect of bias for participants’ ratings of bias in the measurement (33.88 for low bias versus 62.17 for high bias; \(F_{1,127} = 72.52, p < 0.01\)) and management’s forthcomingness (62.61 for low bias versus 55.05 for high bias; \(F_{1,126} = 6.63, p = 0.01\)).

These results indicate a successful manipulation of bias.

**Tests of Hypotheses.** Descriptive (inferential) statistics for participants’ assessments of relevance are presented in Panel A (Panel B) of Table 2. Participants in the low bias condition rated the relevance of the current value information higher than did those in the high bias condition (65.66 for low bias versus 60.29 for high bias); however, this effect is only marginally significant \((p = 0.10, \text{one-tailed})\). Thus, results weakly support the idea that manipulations of bias influence assessed relevance as proposed in Hypothesis 1. No other effects were significant

\(^{14}\) The ANOVA model for assessed bias also shows an insignificant effect of time to sale \((p = 0.79)\) and a significant interaction of bias and time to sale \((p = 0.05)\). This interaction indicates smaller differences in assessed bias between bias conditions in the sell soon condition relative to the no sell condition.
in the ANOVA. Because the manipulation of time to sale did not have the expected effect on participants’ ratings of relevance, we cannot test the directionality of the relationship between reliability and relevance.

[Insert Table 2]

Descriptive and inferential statistics for our valuation dependent measure are also shown in Table 2. The bias manipulation significantly impacted participants’ valuation judgments \((F_{1,113} = 20.89, p < 0.01)\). Specifically, participants in the low bias conditions indicated a larger increase in firm value attributable to the current value measurement (mean of 34.24) than did those in the high bias condition (mean of 4.59). No other effects are significant in the ANOVA.

To test Hypothesis 2, we attempted to fit our hypothesized model to the data, as in Experiment 1. Because of the weak results described above, the model did not provide an appropriate fit for the relationships being tested. Therefore, we are unable to test the idea that preparer bias, an underlying property of reliability, influences both assessed reliability and relevance, which, in turn, separately influence valuation judgments. A possible interpretation of these results is that factors underlying reliability have differential impacts on users’ assessments of relevance. Manipulations of both error (Experiment 1) and bias (Experiment 2) strongly impacted users’ assessments of reliability, but only the former strongly affected assessments of relevance.

5. **Experiment 3**

*Overview*

Experiments 1 and 2 provide evidence supporting our theory that variations in properties underlying reliability influence the assessed relevance of current value information, but that evidence is of varying strength. Additionally, Experiment 1 supports the idea that the variations
in a property underlying reliability (i.e., source competence) influence users’ valuation judgments by two distinct paths—through both assessed reliability and assessed relevance.

In our third experiment, we test whether our theory holds for a third property related to the reliability of a measurement—whether the measurement arises from a market transaction of a comparable item or from a model. We also attempt a stronger manipulation of factors underlying relevance in Experiment 3, in order to allow a convincing test of the directionality of the relationship between relevance and reliability.

Participants, Design, and Variables

Ninety-nine MBA students from a Business Week top-25 school volunteered to participate in Experiment 3 in return for a small fixed payment. The participants had an average of 5.0 years of work experience. Seventy-two percent of participants had previously made investments in common stocks, and 97 percent either had invested or planned to invest in common stocks in the future. Participants had taken an average of 4.7 finance and 3.0 accounting classes. Participants were paid $10 for their participation in this as well as an unrelated study.\(^{15}\)

As in Experiments 1 and 2, participants were asked to assume the role of a potential investor in a fictitious company. All were told that the company had acquired a plot of land for $41 million two years ago that was recently estimated to have a current value of $53 million. Similar to the first two experiments, Experiment 3 makes use of a $2 \times 2$ full-factorial, between-participants design. We manipulate underlying properties of both the reliability of the current value measurement and the relevance of current value information. Specifically, we manipulate the valuation method used in deriving the land’s current value as based on a market price of a comparable asset (e.g., level 2 of SFAS 157 valuation hierarchy) or estimated from a model (e.g.,

\(^{15}\) Participants completed the unrelated study before participating in Experiment 3. Conditions of the other study were counterbalanced with those of our experiments, alleviating the potential for carryover effects. Statistical tests reveal no systematic carryover effects.
level 3). We did not use SFAS 157 terminology in our materials. The two levels correspond to high and low reliability and are referred to as “comparable” and “model,” respectively.16

We manipulate the intended use of the asset as an underlying property of relevance.17 At one level (“sell,” corresponding to high relevance), the land is owned by a real estate development firm that holds the land for development and sale. At the other level (“use,” corresponding to low relevance), the land is owned by a cosmetics manufacturing firm that holds it for purposes of building its own factory. We expect this manipulation will be stronger than that of time to sale in Experiments 1 and 2. Intended use incorporates time to sale (i.e., presumably the real estate developer will sell the land first, making its fair value potentially more timely) and separability of the asset from other operations (i.e., the separate value of land to be developed and sold is likely more relevant for the assigned valuation task than is the fair value of land that is part of productive capacity, because the latter would be valued “in use” rather than separately) (SFAS 157). Figure 3 illustrates the manipulations.

[Insert Figure 3]

After being presented with the information described above, participants made four assessments. The first three response measures are identical to the first three measures captured in Experiments 1 and 2 (i.e., relevance, valuation effect, and reliability). Participants also

---

16 Following the argument of Song et al. (2010), this manipulation is a joint manipulation of bias and potential for error. Specifically, those authors contend that investors’ downward valuation adjustment for reported Level 3 fair values is likely due to two factors: (1) a downward adjustment to the cash flow effect of the item due to potential bias on the part of management, and (2) an upward adjustment to the discount rate applied to the reported amount due to uncertainty or noise in the fair value estimate.

17 The issue of whether intended use should normatively affect relevance judgments was ardently debated during the deliberations of SFAS 157, as verified by Jim Leisenring, then member of the IASB (and liaison to the FASB) and by Robert Lipe, then a faculty fellow at the FASB. For the purposes of this study, however, it is not necessary to establish whether individuals should or should not normatively react to intended use, but only that it will influence assessments of relevance and is unrelated to reliability.
assessed management trustworthiness using a 101-point scale ranging from “not at all trustworthy” (0) to “extremely trustworthy” (100).

Results

Preliminary Tests. As anticipated, reliability was assessed higher by those in the comparable conditions (mean of 64.06) as compared with those in the model conditions (mean of 42.60) \((F_{1,95} = 31.76, p < 0.01)\). Neither intended use \((p = 0.67)\) nor the interaction of intended use and valuation method \((p = 0.65)\) is significant in the ANOVA (not tabulated) for participants’ assessments of reliability. Similarly, an ANOVA model for participants’ ratings of management’s trustworthiness (not tabulated) revealed a significant main effect of valuation method \((F_{1,95} = 13.83, p < 0.01)\). Participants in the comparable conditions rated management trustworthiness significantly higher than did those in the model conditions (58.78 for comparable versus 49.52 for model). Neither intended use \((p = 0.92)\) nor the interaction of intended use and valuation method \((p = 0.27)\) is significant in the model. These results indicate that our manipulation of the valuation method used to derive the asset’s current value successfully influenced participants’ views of the reliability of the asset’s value.

Tests of Hypotheses. Descriptive (inferential) statistics for participants’ assessments of relevance are presented in Panel A (Panel B) of Table 3. The ANOVA reveals a significant main effect of valuation method \((F_{1,95} = 8.44; p < 0.01)\) such that participants in the comparable conditions judged the relevance of current value information (mean of 68.49) significantly higher than did those in the model conditions (mean of 53.94). This result supports Hypothesis 1.\(^{18}\)

\(^{18}\) To determine whether these results hold even if participants are provided with the FASB’s definitions of relevance and reliability, we re-ran Experiment 3 with a separate group of 94 participants from a similar population (MBA students with mean work experience of 8.4 years, average of 3.5 finance and 3.1 accounting classes, 81% had previously invested in common stocks, and 89% planned to do so). We observed that the manipulation of valuation method influenced participants’ assessments of relevance \((F_{1,90} = 7.38; p < 0.01)\) despite the prominent display of FASB’s definitions in the materials. Participants assessed relevance as higher in the comparable conditions (i.e., high reliability) (mean of 75.60) than in the model conditions (i.e., low reliability) (mean of 62.95).
The ANOVA for relevance also reveals a significant main effect of intended use ($F_{1,95} = 4.97; p = 0.03$) and an insignificant interaction of valuation method and intended use ($F_{1,95} = 0.19; p = 0.67$). The significant effect of intended use arises from participants in the sell conditions assessing higher relevance for current value information (mean of 66.68) than did participants in the use conditions (mean of 55.49). This result indicates that we successfully manipulated participants’ views of relevance independent of reliability.

The fact that our manipulation of intended use significantly impacted assessed relevance without impacting assessed reliability provides evidence in support of the unidirectional relationship between relevance and reliability. While factors underlying reliability affect both assessed reliability and assessed relevance, factors underlying relevance influence only assessed relevance. To test further the direct effect of these underlying properties on users’ assessments of reliability and relevance as well as the indirect effect of these properties on users’ valuation decisions through their effect on relevance and reliability, we employ the structural-equations model presented below.

As shown in Table 3, an ANOVA model for participants’ valuation judgments revealed significant main effects of intended use and valuation method. Valuation judgments were more positive when the current value was based on a comparable market transaction (mean of 40.08) versus a model (mean of 19.12) ($F_{1,95} = 8.18, p < 0.01$). They also were more positive when the company intended to develop and sell the asset (mean of 36.32) than when they were intending to use the asset in operations (mean of 22.17) ($F_{1,95} = 3.71, p = 0.06$). There was no evidence of an interaction of these two variables ($p = 0.44$).
For Hypothesis 2, recall that we predict that factors underlying the reliability of a current value measurement will influence valuation via two paths—through users’ assessments of the reliability of the measurement and through their assessments of relevance. Thus, we expect that the valuation method (i.e., comparable versus model) will influence both assessed reliability and relevance and that each will separately and positively influence valuation. We test Hypothesis 2 using a structural-equations model, as presented in Figure 4.

[Insert Figure 4]

We also include in the model the intended use variable, which significantly influenced participants’ relevance judgments. Because we were successful in influencing participants’ judgments of relevance (through the manipulation of intended use) separate from their judgments of reliability (through the manipulation of valuation method), we can test the directional relationship between reliability and relevance. Recall that our theory is that variations in factors underlying reliability will influence relevance judgments, but variations in factors underlying relevance will not influence reliability judgments.

Before exploring the links in the model, we first test the overall fit using various indices. The traditional $\chi^2$ test shows a good fit ($\chi^2 = 1.92, p = 0.59$), as do other standard fit measures.\(^{19}\) Thus, our model provides a good fit for describing the relations in the data.

Path coefficients and statistical significance are presented in Figure 4. The relationship between valuation method, with model coded as 0 and comparable coded as 1, and assessed reliability (Link 1) is positive and significant ($p < 0.01$), consistent with a successful manipulation of factors underlying reliability. Consistent with Hypothesis 1, the relationship between valuation method and assessed relevance (Link 2) is also positive and significant ($p < 0.01$).

---

\(^{19}\) The Comparative Fit Index is 1.00, which is above the generally accepted minimum value of 0.95 (Byrne 2001); the Incremental Fit Index (1.01) is above the recommended minimum of 0.95 (Byrne 2001); and the Root Mean Square Error of Approximation is 0.00, which is below the recommended maximum of 0.01.
Further, both assessed reliability (Link 3, \( p < 0.01 \)) and assessed relevance (Link 4, \( p < 0.01 \)) are positively and significantly associated with valuation, indicating that both influence decision usefulness. Finally, both indirect effects of valuation method on valuation—one through reliability and the other through relevance—are significant \((p < 0.01\text{, product of Z-Scores test})\). These findings support Hypothesis 2.

Turning next to our manipulation of intended use of the asset (with use coded as 0 and sell coded as 1), we see a significant positive relationship with assessed relevance (Link 6, \( p = 0.02 \)), indicating a successful manipulation of factors underlying relevance. Most pertinent to our purposes, the link between the intended use variable and assessed reliability is insignificant (Link 5, \( p = 0.66 \)). This pattern of results supports our assertion that although variations in the reliability of a measurement influence users’ assessments of the item’s relevance, variations in the item’s relevance do not influence users’ assessments of the reliability of its measurement.

**Ruling out a Potential Alternative Explanation.** Our theory predicts that users have difficulty assessing relevance, and so will substitute the reliability attribute for relevance. We believe reliability to be the most likely candidate for attribute substitution because it is a basic property of measurement that prior research indicates individuals assess relatively accurately. Results of three experiments support our theory. However, an alternative possibility is that users substitute decision usefulness for relevance. That is, they may use the label “relevance” to capture the idea of decision usefulness, which includes the effects of both relevance and reliability (FASB 1980, 2010). As explained below, our analysis rules out this possibility.

Recall that our prior analysis shows that participants’ valuation judgments are affected by our manipulations relating to the relevance and reliability of the current value information. If participants used the relevance label to capture decision usefulness, we would observe the
assessed relevance measure capturing *all* effects of factors affecting both relevance and reliability on our valuation dependent measure. However, our analysis does not support this result. Instead, as seen in Figure 4 (Link 3), assessed reliability has a significant effect on valuation above and beyond the impact of assessed relevance on valuation. This finding strongly suggests that our participants did not use relevance as a “stand in” for decision usefulness, but instead viewed decision usefulness as a larger construct that captures their views about both relevance and reliability.20

6. Conclusions

In this paper, we report the results of three experiments in which we investigate the hypothesis that financial statement users’ assessments of the relevance of an economic construct (i.e., fair value) are influenced by factors (i.e., potential for bias and error) underlying the reliability of its measurement. We derive our predictions from psychology theories about attribute substitution and representativeness. Specifically, we find that users’ assessments of the relevance of fair value information are strongly influenced by the competence of the source of the measurement (Experiment 1) and by whether the measurement is based on a market transaction of an analogous item versus a model (Experiment 3). Assessments of relevance are weakly influenced by the expectation of bias in the measurement source (Experiment 2).

Our theory further predicts that the relationship is unidirectional. We find strong support for this idea, as well. In Experiment 3, we find that an underlying property of reliability (i.e., valuation method) influences assessments of the measurement’s reliability; however, it also influences relevance judgments. On the other hand, we find that an underlying property of

---

20 As a further test of this idea, we estimated a model identical to that in Figure 4, except that we omitted Link 3. This model also tests the idea that judged relevance is used as a stand in for decision usefulness. The alternative model is a poor fit to the data, and a nested model test shows that our model fits the data significantly better ($\chi^2 = 8.482, p < 0.01$). These results indicate that our participants did not consider relevance as a substitute for decision usefulness.
relevance (i.e., the intended use of the asset) influences assessed relevance but is not associated with participants’ assessments of reliability. Finally, we demonstrate that properties associated with reliability affect users’ valuation judgments through two paths: 1) their influence on assessed reliability, and 2) their influence on assessed relevance. While we do not have a normative benchmark for the decision weights that should be applied to these properties in making valuation judgments, this evidence that properties related to reliability are “double-counted” is highly suggestive that they are likely to be over-weighted in users’ valuation judgments.

This study has important implications for standard setters, firm management, and financial statement users. First and foremost, our findings suggest that financial statement users are likely to make errors in judgments, including valuations that incorporate assessments of both relevance and reliability. Because these errors are apparently driven by the difficulties of assessing relevance and the resulting use of a heuristic, training may not be sufficient to correct the problem. For example, even individuals with extensive statistical training ignore base rates and rely on representativeness when making intuitive judgments (Tversky and Kahneman 1974, p. 1130).

In addition, our results suggest that a new framework of the FASB and IASB (FASB 2010, ¶QC18; IASB 2010, ¶QC18), in which financial statement preparers and users would assess relevance and reliability sequentially (i.e., consider relevance first and then move to reliability) may not be effective in preventing the problems we document in this paper. That is, our results indicate that users have difficulty considering the relevance of an economic construct separately from factors that influence the reliability of its measurement, and so may be unable to consider relevance first and reliability second. Additional research is warranted on this important issue.
The problems that we document are also likely to interfere with standard setters’ efforts to expand the use of fair value in financial reports. Observers focusing on unreliable fair value measurements have argued that fair values are irrelevant, as a result. We note that this is akin to asserting that the outdoor temperature is not relevant to one’s decision whether to wear a coat simply because the thermometer at hand is unreliable. A better understanding of the underlying issue—users’ inability to assess relevance independently from reliability—indicates that standard setters may wish to focus on improving the reliability of fair value measurements in order to increase acceptance of the idea of the relevance of fair value information.

Finally, our findings that various properties related to the reliability of a measurement (i.e., competence, bias, valuation method) affect users’ relevance judgments differently have important implications for researchers. Our findings allow for a better understanding of the components of information related to reliability that are likely to drive results shown in the archival literature. For instance, Song et al. (2010) find that investors decrease the weight they place on SFAS 157 level 3 fair value measures compared to level 1 and level 2 measures in pricing a firm. The authors indicate that a higher likelihood for both management bias and measurement errors in level 3 measures compared to level 1 and level 2 measures are “likely contributing factors” to their findings. The use of an experiment allows us to confirm what Song et al. (2010) can only presume in their archival research setting. More importantly, our findings suggest that the discounts investors apply to level 3 values may be too large.
REFERENCES

American Bankers Association. 2009. What people are saying about fair value.  
http://www.aba.com/Press+Room/PR_Accounting_FVAQuotes.htm#Steve.


High Competence, Sell Soon Condition

Several years ago, WRSwann Co., a manufacturing company, purchased a plot of land for $41,000,000.

You learn that management intends to sell the plot of land very soon.

The current value of the land is estimated at $53,000,000. This current value number was generated by external consultants.

<table>
<thead>
<tr>
<th>Original Value</th>
<th>Consultants’ Estimate of Current Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$41,000,000</td>
<td>$53,000,000</td>
</tr>
</tbody>
</table>

The consultants generated this number with a model that they created. The model uses the market values for several comparable plots of land. However, no individual plot is completely comparable to the one owned by WRSwann, so the model incorporates additional factors that could affect its value (i.e., location, size, desirability, population growth). The weights that are placed on these additional factors can vary depending on the various assumptions made by the consultants.

Based on your experience and knowledge, you believe the consultants are unbiased, and, based on the accuracy of their work you have noted on other valuations, you believe they are highly competent.
Fig. 2 – This figure summarizes results from Experiment 1. It illustrates that manipulated source competence, a property related to reliability, significantly influences users’ assessments of both the reliability of the measurement and the relevance of current value information. It further demonstrates that users take their assessments of both reliability and relevance into account when determining how much weight to put on the current value information in valuing the firm, indicating that the influence of properties relating to reliability are likely to be double-counted in valuation judgments.
Panel A: Sell, Model Condition

In Year 1, WRSwann Co. purchased a plot of land for $41,000,000. WRSwann Co. is a real estate development firm that routinely purchases, develops, and sells land. The land is ideally suited for development, given its proximity to dense, high-income neighborhoods and transportation.

In Year 3, WRSwann Co. places a current value on the land in the amount of $53,000,000.

This current value number was generated internally by company management. Management based the current value number on an internally developed model (i.e., marked-to-model). There are no comparable plots of land upon which to determine the current value; hence, management used a model to determine the value.

<table>
<thead>
<tr>
<th>Year 1 Value</th>
<th>Year 3 Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$41,000,000</td>
<td>$53,000,000</td>
</tr>
</tbody>
</table>

Panel B: Use, Comparable Condition

In Year 1, WRSwann Co. purchased a plot of land for $41,000,000. WRSwann Co. is a cosmetics manufacturing company that plans to build its factory on the land. The land is ideally suited for the factory, given its proximity to an appropriate workforce and an already established distribution center.

In Year 3, WRSwann Co. places a current value on the land in the amount of $53,000,000.

This current value number was generated by an external valuation firm. This external valuation firm based the current value number on a comparable plot of land (i.e., adjacent to the one purchased by WRSwann, same size, and same location desirability). This comparable plot of land was recently sold.

<table>
<thead>
<tr>
<th>Year 1 Value</th>
<th>Year 3 Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$41,000,000</td>
<td>$53,000,000</td>
</tr>
</tbody>
</table>
FIGURE 4
Structural-Equations Model, Experiment 3

Fig. 4 - This figure summarizes results from Experiment 3. It illustrates that manipulated valuation method (i.e., comparable transaction versus model), a property related to reliability, significantly influences users’ assessments of both the reliability of the measurement and the relevance of current value information. It demonstrates the directionality of this “confusion” between reliability and relevance by showing that the assets’ intended use (use as a factory site vs. held for sale), a determinant of relevance, significantly influences assessed relevance but does not influence assessed reliability. It further demonstrates that users take their assessments of both reliability and relevance into account when determining how much weight to put on the current value information in valuing the firm, indicating that the influence of properties relating to reliability are likely to be double-counted in valuation judgments.
Table 1 - Participants assessed the relevance of the current value information to their overall evaluation of the firm. The competence of the external consultants determining the measurement was varied at two levels, high and low. Management’s plans to sell the asset were varied at two levels, sell soon versus no sell. Participants provided relevance assessments on a 101-point scale, with 0 labeled “not at all relevant” and 100 labeled “extremely relevant.” Participants also assessed whether the current value information increases, decreases, or neither increases nor decreases the value they place on the firm. Participants who indicated a change in value also indicated the magnitude of the change on a 101-point scale, ranging from “slightly increases (decreases) how much I value the company” (0) to “greatly increases (decreases) how much I value the company” (100). Valuation effect is calculated by coding the magnitude of change as a positive amount for participants who indicated increases in value, a negative amount for those who indicated decreases in value, and zero for participants who indicated no change in value.

<table>
<thead>
<tr>
<th></th>
<th>Assessed Relevance Dependent Variable</th>
<th>Valuation Effect Dependent Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Preparer Competence</td>
<td>Preparer Competence</td>
</tr>
<tr>
<td>Time to Sale</td>
<td>Low (24.40)</td>
<td>High (21.78)</td>
</tr>
<tr>
<td></td>
<td>n = 30</td>
<td>n = 33</td>
</tr>
<tr>
<td>No Sell</td>
<td>47.47 (24.40)</td>
<td>57.29 (24.74)</td>
</tr>
<tr>
<td></td>
<td>n = 30</td>
<td>n = 63</td>
</tr>
<tr>
<td>Sell Soon</td>
<td>54.94 (22.92)</td>
<td>59.86 (26.66)</td>
</tr>
<tr>
<td></td>
<td>n = 34</td>
<td>n = 66</td>
</tr>
<tr>
<td></td>
<td>16.12 (35.20)</td>
<td>34.24 (33.42)</td>
</tr>
<tr>
<td></td>
<td>n = 64</td>
<td>n = 59</td>
</tr>
<tr>
<td>Column Means</td>
<td>51.44 (23.73)</td>
<td>65.66 (25.73)</td>
</tr>
<tr>
<td></td>
<td>n = 64</td>
<td>n = 65</td>
</tr>
</tbody>
</table>

Panel B: Analysis of variance

<table>
<thead>
<tr>
<th>Source</th>
<th>Assessed Relevance</th>
<th>Valuation Effect</th>
<th>Two-tailed p-value</th>
<th>Two-tailed p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time to Sale</td>
<td>1 325</td>
<td>F = 0.53</td>
<td>0.47</td>
<td>1 1060</td>
</tr>
<tr>
<td>Preparer Competence</td>
<td>1 6718</td>
<td>F = 10.92</td>
<td>&lt; 0.01</td>
<td>1 9944</td>
</tr>
<tr>
<td>Time to Sale × Preparer Competence</td>
<td>1 594</td>
<td>F = 0.97</td>
<td>0.33</td>
<td>1 1942</td>
</tr>
<tr>
<td>Residual</td>
<td>125 615</td>
<td>114 1176</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# TABLE 2
## Relevance Assessments and Valuation Judgments, Experiment 2

### Panel A: Mean judgments (standard deviation)

<table>
<thead>
<tr>
<th></th>
<th>Assessed Relevance Dependent Variable</th>
<th>Valuation Effect Dependent Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Preparer Bias</td>
<td>Preparer Bias</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td><strong>Time to Sale</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Sell</td>
<td>66.21 (21.78)</td>
<td>59.28 (24.04)</td>
</tr>
<tr>
<td></td>
<td>n = 33</td>
<td>n = 36</td>
</tr>
<tr>
<td>Sell Soon</td>
<td>65.09 (29.61)</td>
<td>61.50 (15.27)</td>
</tr>
<tr>
<td></td>
<td>n = 32</td>
<td>n = 30</td>
</tr>
<tr>
<td><strong>Column Means</strong></td>
<td>65.66 (25.73)</td>
<td>60.29 (20.40)</td>
</tr>
<tr>
<td></td>
<td>n = 65</td>
<td>n = 66</td>
</tr>
</tbody>
</table>

### Panel B: Analysis of variance

<table>
<thead>
<tr>
<th>Source</th>
<th>Assessed Relevance</th>
<th>Valuation Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>df</td>
<td>MS</td>
</tr>
<tr>
<td>Time to Sale</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>Preparer Bias</td>
<td>1</td>
<td>903</td>
</tr>
<tr>
<td>Time to Sale × Preparer Bias</td>
<td>1</td>
<td>91</td>
</tr>
<tr>
<td>Residual</td>
<td>127</td>
<td>545</td>
</tr>
</tbody>
</table>

Table 2 – Participants assessed the relevance of the current value information to their overall evaluation of the firm. The bias of the source of the current value was varied as either high or low. Management’s plans to sell the asset were also varied at two levels, sell soon versus no sell. The dependent measures are described in the notes to Table 1.
### TABLE 3
Relevance Assessments and Valuation Judgments, Experiment 3

Panel A: Mean judgments (standard deviation)

<table>
<thead>
<tr>
<th>Intended Use</th>
<th>Assessed Relevance Dependent Variable</th>
<th>Valuation Effect Dependent Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model</td>
<td>Comparable</td>
</tr>
<tr>
<td><strong>Use</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use</td>
<td>49.48 (24.13)</td>
<td>61.75 (25.49)</td>
</tr>
<tr>
<td>n = 25</td>
<td>n = 24</td>
<td>n = 49</td>
</tr>
<tr>
<td>Sell</td>
<td>58.40 (24.84)</td>
<td>74.96 (24.29)</td>
</tr>
<tr>
<td>n = 25</td>
<td>n = 25</td>
<td>n = 50</td>
</tr>
<tr>
<td>Column Means</td>
<td>53.94 (24.65)</td>
<td>68.49 (25.51)</td>
</tr>
<tr>
<td>n = 50</td>
<td>n = 49</td>
<td></td>
</tr>
</tbody>
</table>

Panel B: Analysis of variance

<table>
<thead>
<tr>
<th>Source</th>
<th>Assessed Relevance</th>
<th>Valuation Effect</th>
<th>Two-tailed p-value</th>
<th>Two-tailed p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intended Use</strong></td>
<td>df</td>
<td>MS</td>
<td>Statistic</td>
<td>df</td>
</tr>
<tr>
<td>Intended Use</td>
<td>1</td>
<td>3029</td>
<td>$F = 4.97$</td>
<td>0.03</td>
</tr>
<tr>
<td>Valuation Method</td>
<td>1</td>
<td>5141</td>
<td>$F = 8.44$</td>
<td>0.01</td>
</tr>
<tr>
<td>Intended Use × Valuation Method</td>
<td>1</td>
<td>113</td>
<td>$F = 0.19$</td>
<td>0.67</td>
</tr>
<tr>
<td>Residual</td>
<td>95</td>
<td>609</td>
<td></td>
<td>95</td>
</tr>
</tbody>
</table>

Table 3 – Participants assessed the relevance of the current value information to their overall evaluation of the firm. We manipulated whether the current value was derived from a market transaction for a similar asset (e.g., level 2 of SFAS 157 valuation hierarchy) or from a model (e.g., level 3). We also manipulated the intended use of the asset as either “use” (the land is held for purposes of building the owner's factory) or “sell” (the land is held for development and sale). The dependent measures are described in the notes to Table 1.