When the use of positive language backfires : The joint effect of tone, readability, and investor sophistication on earnings judgments

Tan, Hun-Tong; Ying Wang, Elaine; Zhou, Bo

2014

When the Use of Positive Language Backfires:

The Joint Effect of Tone, Readability, and Investor

Sophistication on Earnings Judgments

Hun-Tong Tan
Nanyang Business School
Nanyang Technological University
ahttan@ntu.edu.sg

Elaine Ying Wang
Isenberg School of Business
University of Massachusetts Amherst
ewang@isenberg.umass.edu

Bo Zhou
School of Accountancy
Shanghai University of Finance and Economics
zhou.bo@mail.shufe.edu.cn

Dec 11, 2013

When the Use of Positive Language Backfires:

The Joint Effect of Tone, Readability, and Investor Sophistication on Earnings Judgments

ABSTRACT

Recent studies document that market participants react positively to the positive language sentiment or tone embedded in financial disclosures, and that investors' reactions to negative news are more muted with poor disclosure readability. However, while language sentiment and readability co-occur in practice, their joint effects remain largely unexplored. In an experiment with MBA students as participants, we investigate how the effect of language sentiment varies with readability and investor sophistication level. We find that language sentiment influences investors' judgments when readability is low, but not when readability is high. Specifically, when readability is low, disclosures couched in positive language lead to higher earnings judgments for less sophisticated investors, but lower earnings judgments for more sophisticated investors. These findings show that the main effects of readability and language sentiment documented in prior studies have boundary effects, and may reverse when both variables are jointly considered along with investor sophistication.

Keywords: *tone; language sentiment; readability; qualitative disclosures; investor sophistication; information asymmetry*

1. Introduction

Prior research shows that managers' disclosures of firm performance vary in terms of language sentiment or tone, and that such variation in tone is positively associated with short-term market reaction (Davis, Piger and Sedor [2012], Feldman et al. [2010], Huang, Teoh, and Zhang [2013]). However, because the tone in management disclosures can contain information content, it is unclear whether the positive reaction observed in these studies is due to additional information content (e.g., Davis, Piger and Sedor [2012]), or a language sentiment effect associated with the mere use of linguistic terms (e.g., positive versus neutral words) to describe an identical hard number or quantitative disclosure (e.g., Huang, Teoh, and Zhang [2013]).[1] Investigating whether a pure language sentiment effect exists is important because it suggests that language sentiment could then be used by managers to systematically influence investors' judgments, holding information content constant.

In this study, we investigate how the language sentiment effect is conditional on readability and investor sophistication. The two moderators—readability and investor sophistication—have not been examined in the language sentiment literature, but both are ubiquitous features in the management disclosure environment. For instance, prior studies document systematic differences in the readability of management disclosures (Li [2008]) and that investors' reactions to news content are more subdued with lower readability (Miller [2010], Rennekamp [2012], Tan, Wang, and Zhou [2012], You and Zhang [2009]). Most pertinent to our study, managers contemporaneously vary the use of *both* language sentiment and readability to obfuscate unsatisfactory performance (Henry [2008], Rogers, Buskirk and Zechman [2011]). However, the existing language sentiment and/or readability literature does not inform us on how the language sentiment effect is moderated by readability. Following the current literature, if low readability obfuscates readers' understanding, it may be that low readability also dampens the effect of language sentiment; alternatively, it may be that low

---

[1] For example, Larcker and Tayan [2010] cite the example of Erin Callan, former CFO of Lehman Brothers, who used language that was generic and excessively positive to obscure the company's deteriorating financial position. In a conference call just months before Lehman's collapse, she used the word "great" 14 times, "strong" 24 times, and "incredibly" eight times.

readability magnifies the effect of language sentiment since the former can obfuscate evidence inconsistent with the language sentiment.

Further, while both more and less sophisticated investors are recipients of these management disclosures that vary in language sentiment and readability, it is unknown whether language sentiment effects (as moderated by readability) are contingent on investor sophistication. It may be that this effect is pervasive, such that more sophisticated investors, just like less sophisticated investors, are affected but perhaps to a smaller extent. On the other hand, more sophisticated investors may see through the strategic nature of management's use of positive language and low readability, and "punish" management accordingly by discounting such disclosures. In either case, there will be disagreement among investors, which can increase firms' information asymmetry and cost of capital (Bloomfield and Fischer [2011]). Furthermore, in cases where more sophisticated investors are "smarter" than less sophisticated investors, more sophisticated investors can sometimes drive biased market prices to unbiased levels (Bloomfield, Libby and Nelson [1996]), even though their disciplining role in eliminating market inefficiencies may not always be sufficient (Libby, Bloomfield, Nelson [2002]).

Examining the joint effect of language sentiment, readability, and investor sophistication is important because, as we discuss later, these variables do not necessarily have additive effects on investor judgment. Rather, the effect of language sentiment on investor judgment is conditional on readability (and vice versa), with the directional effects varying depending on investor sophistication level. More specifically, the directional effects of language sentiment as documented in the current literature may actually be muted or even reversed once these interactive effects are considered.

Following psychology research, we consider the language sentiment effect to be an example of attribute framing (Kuvaas and Selart [2004], Levin, Schneider, and Gaeth [1998]). Attribute framing refers to an effect where people's evaluations of two otherwise identical items (e.g., disclosures with identical quantitative/hard

information in our context) will be more positive or negative depending on whether the items are described in a positive or a negative manner (e.g., positive versus neutral linguistic terms in our context). We employ the dual processing (or process) model of cognition (Chaiken [1980], Chaiken, Liberman, and Eagly [1989], Petty and Cacioppo [1986]), a widely accepted model of human information processing in psychology, to develop our theoretical predictions. The dual processing model of cognition posits that people's reasoning can be described as a function of both an intuitive heuristic based system and an analytical deliberative based system. Specifically, it posits how contextual/heuristic cues (i.e., attribute framing in our context) influence people's judgments depends on target feature (i.e., disclosure readability in our context) and judge feature (i.e., investor sophistication in our context).

Building on the dual processing model, we propose that readability and investor sophistication level can jointly moderate the impact of language sentiment on investor judgment. Specifically, we predict that the effect of language sentiment on investor judgment is conditional on readability. When readability is high, the substance of a disclosure can be easily understood, and the language used to frame the message becomes less relevant. Hence, language sentiment is less likely to have an impact on investor judgment, irrespective of investor sophistication. Conversely, when readability is low, language sentiment can have a marked effect on investor judgment, with the directional effect varying depending on investor sophistication level. Less sophisticated investors' earnings judgments are susceptible to positive framing effects arising from positive language used, even though the positive language may be inconsistent with some aspects of the management disclosures that have negative implications for the company's earnings prospects. In contrast, more sophisticated investors are more likely to consider the use of positive language to be less credible, causing them to be more negative when positive (as opposed to neutral) language is used.

We conduct a 2 (positive vs. neutral language) × 2 (high vs. low readability) × 2 (less vs. more sophisticated investors) between-participants experiment with MBA

students to examine our research questions. An experimental method is an appropriate approach here because it enables us to vary the linguistic features while holding constant the quantitative information content. The use of archival data entails difficulty in isolating linguistic features from the information content of the disclosures, an endogenity issue that Li [2010] highlights. We choose earnings release as our experiment context, and employ a context where management presents a seemingly satisfying quarterly performance in the first two paragraphs but key financial highlights in the third paragraph portray an unsatisfying quarterly performance. We manipulate language sentiment in the first two paragraphs of the earnings release by using either positive evaluative words or neutral words to describe an identical set of quantitative financial results. We also manipulate readability in the last two paragraphs of the earnings release by using either the plain English principles issued by SEC [1998] (e.g., simple words, short sentences, organized structures with table/bullet presentation) or the converse of these principles (e.g., complex words, long sentences, mixed structures without table/bullet presentation). Finally, we measure investor sophistication based on participants' demographic information, such as the number of accounting and finance courses the MBA students have taken and their frequency of reading annual reports and earnings releases.

Consistent with our hypotheses, we find that, when readability is high, language sentiment does not affect MBA student participants' earnings judgments, irrespective of their sophistication level. However, when readability is low, the use of positive as opposed to neutral language positively influences less sophisticated participants' earnings judgments. In addition, we conduct structural equation modeling (SEM) to analyze the underlying mechanism through which this effect occurs. We find that positive as opposed to neutral language causes less sophisticated participants to have more positive initial impressions of the firm's current quarter performance (relative to prior management guidance), and also poorer understanding of the firm's unfavorable trend performance (current quarter performance compared with year-ago quarter performance). This impaired understanding in turn leads to over-optimistic earnings

4

judgments. In contrast, the obfuscation strategy of using positive and difficult-to-read language does not work when participants are more sophisticated. In particular, when readability is low, the use of positive language has a negative influence on more sophisticated participants' earnings judgments. Our SEM results show that more sophisticated participants consider positive language written in a difficult-to-read manner to be less credible, causing them to make lower earnings judgments. In addition, we find that information asymmetry (measured by the difference between more and less sophisticated investors' earnings judgments) is larger when readability is low than when it is high, an effect that is magnified when the disclosures are couched in positive language.

Our findings are closely related to the literature on disclosure tone, and provide insights into the non-informational role of tone, namely language sentiment, on investor judgment. It is necessary to distinguish between the tone effect arising from language sentiment and that associated with informational difference. If tone effects arise from informational differences, tone effects are rational and the market should pay attention to the tone embedded in management disclosures (Feldman et al. [2010]). However, if the market reacts merely to positive language absent information content, tone effects would be considered an over-reaction to language sentiment, and one would expect stock prices to potentially reverse in the future (Huang, Teoh, and Zhang [2013], Tetlock [2007]). As a result, profitable trading strategies could be developed to exploit this over-reaction of the market. We provide direct evidence that purely linguistic features (i.e., the positive linguistic terms used rather than positive information) can influence investors' judgments.

We also add to the theory on the language sentiment effect by positing and demonstrating that the effect of language sentiment is more likely when readability is low than when it is high. This finding suggests that a remedy to de-bias the language sentiment effect is to make disclosures more readable. As we show in our results, less sophisticated investors are susceptible to language sentiment only when the disclosure is less readable but not when it is more readable. Thus, the remedy may not

necessarily lie in the form of prescriptions relating to management's strategic exploitation of language sentiment effects in their disclosures—precise guidelines are likely difficult in this instance. Instead, the remedy has already been set in place by the SEC—the SEC's call and guidelines for improved readability in management disclosures (SEC [1998]), if implemented by firms, may serve to reduce the influence of language sentiment.

Our results also have implications for managers when they consider the strategic use of language sentiment and low readability in their disclosures. Specifically, managers should be aware of the negative impact of using difficult-to-read plus positive language disclosures to skew investors' judgments. As documented in our study, sophisticated market participants can see through such strategic disclosure behavior, which in turn results in discounted disclosure credibility and impaired earnings judgments.

Our investigation of language sentiment effects complements the recent study by Hales, Kuang, and Venkataraman [2011], which examines the use of vivid versus pallid language, holding constant the language sentiment (positive in their Experiment 1, and negative in their Experiment 2). In contrast to their study, our focus is on language sentiment (positive versus neutral) and its interaction with readability and investor sophistication level, issues not examined in Hales, Kuang, and Venkataraman [2011]. Furthermore, we show that investor sophistication and other linguistic features (i.e., readability and language sentiment) can also affect investors' judgments towards pallid language.

The next section reviews related literature and develops our hypotheses. Section 3 describes the research design and experiment procedure. Section 4 presents the results. Section 5 concludes.

2. Theory and Hypotheses

2.1 The Role of Positive Language on Investor Judgment

Managers have a tendency to use positive language to describe their firms' performance and prospects (Henry [2008]). The implicit premise is that describing an

identical quantitative performance with positive (versus neutral) language can lead to systematically more favorable impressions. For instance, Macgregor et al. [2000] find that words with positive evaluative meanings can trigger favorable images of the expressed objects in people's minds and further influence their subsequent judgments. Similar findings are documented in the consumer behavior literature. For example, pleasant music in a commercial can lead to a favorable image of the product, even though the music is objectively irrelevant to the merits of the product (Gorn [1982]).

These findings are related to attribute framing phenomenon documented in psychology literature (Levin, Schneider, and Gaeth [1998], Tversky and Kahneman [1986]). Specifically, attribute framing involves influencing people's judgments of an object or event by describing it in a positive or a negative manner, while holding its objective value constant. The typical finding in this line of literature is that an object or an event is evaluated more favorably when it is presented in a positive frame relative to a negative frame. As an example, consumers' evaluations are more favorable toward a beef product labeled "75% lean" than one labeled "25% fat" (Levin and Gaeth [1988]). The theoretical explanation offered for attribute framing is that information is encoded in relation to its descriptive valence. While positive labeling of an attribute leads to an encoding of the information that evokes favorable associations in memory, negative labeling of the same attribute causes an encoding that evokes unfavorable associations.

Attribute framing has been found in a variety of cognitive tasks (e.g., attention, learning, memory, and associations) and settings (e.g., marketing/advertisement, political science, and health care) (Kuvaas and Selart [2004], Levin, Schneider, and Gaeth [1998]). It has been shown to occur even in field settings with repeated interactions (Hossain and List [2012]). Attribute framing can influence subsequent evaluations and judgments, and the extent to which this occurs depends on how individuals process the information.

According to the dual-process model, there are two different processing routes, which can lead to different outcomes. One route involves a relatively controlled and

analytic processing style, in which the ultimate judgment is based on careful and effortful evaluation of available information. The other route is relatively automatic and holistic such that the ultimate judgment is not based on extensive processing of message but on some superficial cues or simple "rules of thumb" (Chaiken [1980], Chaiken, Liberman, and Eagly [1989], Petty and Cacioppo [1986]). Individuals who process information in a holistic style rely on contextual cues and make inferences about the issue without detailed scrutiny of the material. Consequently, this type of processing should be especially sensitive to contextual cues, such as how the problem is framed (McElroy and Seta [2003]). In contrast, contextual cues are less likely to influence individuals who process information in a more analytic style.

The dual-process model proposes that the target feature (e.g., complexity of the task) and the judge feature (e.g., cognitive capacity) can influence the interplay of the two processing routes (Chaiken, Liberman, and Eagly [1989], Petty and Cacioppo [1986]). When the task is complicated and requires a great amount of resources to complete, people are more likely to engage in heuristic processing. In addition, the cognitive capacity of the judge also plays an important role, with heuristic rather than substantive processing more likely to be used when the judge's processing capacity is in some way limited. In the current study, our target feature of interest is the readability of an earnings release, and our judge feature of interest is the sophistication level of investors. Below, we further discuss how readability and investor sophistication affect the impact of language sentiment on investor judgment.

2.2 Less Sophisticated Investors

We posit that language sentiment positively influences less sophisticated investors' earnings judgments. The positive language in the earnings release can result in positive framing effects and thus favorable impressions, which serve as heuristic cues for people's judgments. When people lack the ability to deeply process the issues being considered, they use their impressions of the message as the basis for their judgments (Petty et al. [1993]). Since less sophisticated investors may not have sufficient ability to process the information embedded in an earnings release, we

expect that they will be susceptible to positive framing effects (i.e., make higher earnings judgments when reading disclosures with positive language compared to those with neutral language).

Because low readability increases investors' processing difficulty, we propose that the framing effects arising from positive language likely influence less sophisticated investors to a greater extent when a financial disclosure is difficult to read than when it is easy to read. Specifically, when readability is high, less sophisticated investors can better understand the underlying information and are less likely to be influenced by managers' use of positive language. In contrast, low readability increases the information processing difficulty. As a result, less sophisticated investors are more likely to rely on heuristic cues such as language sentiment to make their judgments. We state our hypotheses regarding less sophisticated investors' earnings judgments below:

H1: Less sophisticated investors' evaluations of firm's future earnings performance are more favorable when the earnings release is positively written than when it is neutrally written, and the difference is larger when the earnings release is more difficult to read.

2.3 More Sophisticated Investors

Psychology research indicates that while heuristic processing is more likely when decision makers have lower processing ability, systematic processing is more likely when they have higher processing ability (Chaiken [1980], Chaiken, Liberman, and Eagly [1989], Petty and Cacioppo [1986]). Since more sophisticated investors have the requisite knowledge and ability to process management disclosures, we expect that more sophisticated investors will likely process the information in a more analytic and systematic style. This systematic processing enables readers to better understand the conveyed message, and immunizes them from being influenced by contextual features such as language sentiment (McElroy and Seta [2003]). As a result, more sophisticated investors are less likely to suffer from attribute framing effects. This suggests that the favorable impact of positive language and low readability on

investors' earnings judgments that we discussed in H1 will be weakened for more sophisticated investors. In fact, as we discuss below, more sophisticated investors may be more skeptical of positive language used in conjunction with low-readability disclosures.

In the context of management disclosures, research shows that positive words relative to negative words have less predictability for future returns (Engelberg [2008]), and that the use of positive language is incentive-consistent and less credible (Han and Tan [2010], Hutton, Miller, and Skinner [2003]). In particular, people's attempts to undo biases are likely to be more successful when these people have the ability than when they do not (Wegener and Petty [1995]). Accordingly, more sophisticated investors are also more likely to have a better understanding of management disclosures and see through management's self-serving intentions. This suggests that more sophisticated investors are more likely to penalize unsupported positive language and reward neutral evaluative language. We also expect this punitive effect of positive language to be more pronounced when readability is low, because low readability obfuscates the information conveyed in the positive-language disclosure and leads to more suspicion about the credibility of the disclosed positive-language information. We state our hypothesis regarding more sophisticated investors' earnings judgments below:

H2: More sophisticated investors' evaluations of firm's future earnings performance are more favorable when the earnings release is neutrally written than when it is positively written, and the difference is larger when the earnings release is more difficult to read.

3. Experiment

3.1 Participants

One hundred and forty two Masters of Business Administration (MBA) students (average work experience of 5.97 years) from two major universities in Singapore participate in our experiment. On average, participants have taken 2.34 accounting courses and 2.87 finance courses. About 95% percent of them have previously read

annual reports and/or earnings releases. According to Elliott et al. [2007], actual retail investors on average have taken 1.9 accounting courses, 1.6 finance courses, and 97% of them have previously evaluated financial statements. Hence, our MBA participants are appropriate proxies for retail investors as their profiles on key attributes are similar to those in Elliott et al. [2007].

3.2 Investor Sophistication Level

We use a $2 \times 2 \times 2$ (language sentiment by readability by investor sophistication level) between-participants experiment to test our hypotheses. We measure investor sophistication level based on demographic information, such as the number of accounting and finance courses participants have taken, their frequency of reading annual reports, and their frequency of reading earnings releases.[2] We calculate the sophistication score of participants by summing up the standardized values of the responses listed above with equal weight. The sophistication score ranges from -3.08 to 11.15, with a median of -0.60. We split participants into more sophisticated investors and less sophisticated investors based on the median sophistication score. The mean sophistication scores for less and more sophisticated investors are -1.45 and 1.55, respectively. More specifically, with respect to less sophisticated investors, the average number of accounting (finance) courses taken is 1.15 (1.49), and the average frequency of reading annual reports (earnings releases) is 1.97 (2.11), on a scale of 0 (never) to 10 (with high frequency). With respect to more sophisticated investors, the average number of accounting (finance) courses taken is 3.53 (4.25), and the average

---

[2] We include the frequency of reading annual reports and the frequency of reading earnings releases in our sophistication score because the frequency of analyzing financial statements is one attribute that distinguishes early MBA students (less sophisticated) from select MBA students (more sophisticated), as shown in Elliott et al. [2007]. We exclude work experience because Elliott et al. [2007] document that this attribute does not differentiate performance of MBA students and retail investors, suggesting that work experience is not a deterministic factor of investor sophistication level. We also conduct a factor analysis (i.e., principal components analysis with varimax rotation) on the demographic information we collected. The results show that two components are extracted, with frequency of reading annual reports and earnings releases loading heavily onto Component 1 (factor loadings are 0.96 and 0.96 for Component 1, and 0.03 and 0.09 for Component 2, respectively), and accounting and finance courses loading heavily onto Component 2 (factor loadings are 0.88 and 0.88 for Component 2, and 0.04 and 0.07 for Component 1, respectively). The factor loadings for work experience are low at 0.41 for Component 1 and 0.13 for Component 2. As the factor loadings are smaller than the cut-off benchmark of 0.70 for inclusion as a factor component (Carmines and Zeller [1979]), we exclude work experience when calculating our investor sophistication measure.

frequency of reading annual reports (earnings releases) is 5.13 (4.75).

The study by Elliott et al. [2007] provides further support for the validity of our investor sophistication measures. Elliott et al. [2007] compare two groups of MBA students' performance with that of retail investors, and find that early MBA students (students in the early stage of the MBA program) perform worse than retail investors when task complexity is relatively high, while select MBA students (second-year MBA students) perform slightly better than retail investors. These results suggest that early and select MBA students represent two different groups of investors in terms of sophistication level. In their study, early (select) MBA students have taken 1.8 (3.5) accounting and 1.0 (2.9) finance courses, have 5.2 (5.9) years of work experience, and 17% (7%) of them have never evaluated financial statements before. In our study, the less (more) sophisticated participants have taken 1.15 (3.53) accounting and 1.49 (4.25) finance courses, have 5.3 (6.6) years of work experience, and 9.9% (none) of them have never read annual reports or earnings releases before. Hence, our less (more) sophisticated participants are comparable to the early (select) MBA students participated in Elliott et al. [2007], and we expect that our participants can well represent investors with different sophistication levels.

3.3 Manipulation of Readability and Language Sentiment

When managers issue earnings press releases, they usually begin with broad comments on the firm's overall earnings performance, followed by more detailed discussion of specific earnings components (Henry [2008]). Further, variations in management tone are often found in the section of the earnings releases where managers make broad comments about the firm's performance (Feldman et al. [2009]), while variations in readability are often found in the section where detailed financial information is presented (Davis and Tama-Sweet [2012]). Hence, we manipulate language sentiment (positive vs. neutral) in the first two paragraphs of the earnings release, and readability (high vs. low) in the last two paragraphs of the earnings release. We restrict our manipulations of language sentiment and readability to distinct (rather than identical) parts of the earnings release to avoid the possibility that

manipulation of language sentiment for the same information inadvertently also varies its readability (and vice versa).

The language sentiment manipulation relates to management's comments on the current quarter earnings performance, and we hold the actual earnings numbers and associated information constant while varying only the language used to describe this part of the earnings release. We develop the positive language manipulation based on earnings releases issued by publicly listed companies, and then modify them to create the neutral language manipulation. For example, in the positive language manipulation, management employs positive language such as "strong," "significantly higher," "far above," "pleased," or "impressed." In the neutral condition, these adjectives are either deleted or replaced with more neutral language. As an illustration, in the positive language condition, the news release indicates: "The Company reported strong second quarter results that achieved the company's overall expectations. Earnings per share were $1.08, significantly higher than the low end of company's guidance range of $1.02 to $1.09." In contrast, the neutral earnings release indicates: "The Company reported second quarter results that were roughly in line with the company's overall expectations. Earnings per share were $1.08, within the company's guidance range of $1.02 to $1.09." Appendix 1 provides details of our language sentiment manipulation.

We manipulate readability of the earnings release in the last two paragraphs of the earnings release. We first create the low readability version by referring to actual earnings releases issued by listed firms that employ complex words, long sentences, and complex presentation layouts (without table/bullet presentations), all of which are features that contradict the plain English writing principles issued by the SEC. We then rewrite the release to create a high readability version following plain English writing principles.[3] Specifically, in the high readability condition, the earnings

---

[3] The plain English writing principles include: (i) short sentences; (ii) definite, concrete, everyday words; (iii) active voice; (iv) tabular presentation or bullet lists for complex material whenever possible; (v) no legal jargon or highly technical business terms; and (vi) no multiple negatives (SEC [1998]). Our readability manipulations only involve principles (i), (ii), and (iv). We hold the other principles constant across conditions.

release is written with simple words and short sentences, along with table/bullet presentations (SEC [1998], Loughran and McDonald [2010]). The latter feature is less frequently used by firms (relative to the use of simple words and short sentences), but is recommended by both the SEC and linguistic research as an approach that improves readability (SEC [1998], Cutts [1996]). Appendix 2 provides details of our readability manipulation.

3.4 Procedure

We ask participants to assume the role of an investor. Participants first read some background information and financial data about a listed company (W&Z), and then read an earnings release that consists of four paragraphs. The first two paragraphs contain management's assessment of the current quarter performance and discussion of the firm's future prospects. The language in these two paragraphs is couched either in positive or neutral language. The next two paragraphs of the earnings release present the key financial highlights and management earnings guidance for the next quarter, which vary in readability.

The key financial highlights section presents seven performance indicators in comparison to those in the year-ago quarter (see details in Appendix 2). Among the seven performance indicators, only two performance indicators (e.g., gross margin and earnings from discontinued operations) increased from the year-ago quarter. The other five performance indicators, including EPS and earnings from continuing operations (which are predictive of future performance; see Fairfield, Sweeney and Yohn [1996]), all decreased from the year-ago quarter. Hence, if participants fully understand the information in the key financial highlights section, they should infer that the firm's current-quarter financial performance is poorer relative to that in the year-ago quarter. In the future earnings guidance section, management upwardly revised the guidance on core sales growth, net sales, foreign exchange impact, and earnings per share.

After reading the earnings release, participants make their judgments regarding the future performance of the target company's earnings. We also ask participants to

assess the believability and credibility of the earnings release. Once participants complete the questions on the main dependent variables, we ask them to insert all the materials into Envelope A and open Envelope B.

Envelope B contains debriefing questions, as well as manipulation check and demographic questions. We assess whether positive language influences participants' initial impressions of the firm's performance by asking participants to describe management's assessment/comments/confidence pertaining to the firm's current quarter performance. We then assess participants' understanding of the firm's performance compared with the year-ago quarter. We also measure participants' assessments of other attributes of the earnings release, such as its reporting clarity. The final section elicits participants' responses to the manipulation check and demographic questions.

4. Results

4.1 Manipulation Checks

To assess the effectiveness of our manipulations, we ask participants to evaluate several aspects of the earnings release that they read. In the post-experimental questionnaire, participants first rate the language sentiment in the first two paragraphs of W&Z's earnings release. Specifically, they evaluate President Bill Sloan's manner of speaking, the management's attitude toward W&Z's overall performance, and the tone of management's communication on a 13-point Likert scale, with endpoints 0 = "Not at all Favorable" and 12 = "Extremely Favorable." The mean ratings in the positive language condition (neutral language condition) are 7.90 (5.48)/7.96 (5.92) /8.30 (5.75) for President Bill Sloan's manner of speaking / the management's attitude toward W&Z's overall performance / the tone of management's communication, respectively. Further, the mean ratings in the positive language condition are all significantly higher than those in the neutral language condition, $p < 0.01$,[4] suggesting that participants perceive the language in the positive language condition

---

[4] All p-values are two-tailed unless specifically specified.

to be more favorable than that in the neutral language condition.[5]

Participants then rate the readability of the message in the last two paragraphs of W&Z's earnings release (i.e., key financial highlights and guidance for the next quarter). Specifically, they evaluate the extent to which they think the message is difficult to read / difficult to understand / difficult to process on a 13-point Likert scale, with endpoints 0 = "Not at all Difficult" and 12 = "Extremely Difficult." The mean ratings in the high readability condition (low readability condition) are 4.71 (6.29)/4.46 (5.93)/5.49 (6.30) for the difficult to read/ difficult to understand/ difficult to process measures, respectively. The mean ratings in the high readability condition are significantly lower than the mean ratings in the low readability condition ($p < 0.01$ for the difficult to read / difficult to understand measures; one-tailed $p = 0.04$ for the difficult to process measure). These results generally suggest that participants perceive the message in the low readability condition to be more difficult to read/understand/process than that in the high readability condition, indicating that our readability manipulation is successful.[6]

4.2 Hypothesis Testing

4.2.1 All Investors

We measure participants' evaluations of W&Z's future earnings performance by asking them to indicate the extent they agree that the firm's earnings performance will be strong in the near future on an 11-point scale, with "-5" labeled "strongly disagree," "5" labeled "strongly agree," and "0" labeled "neutral." The descriptive statistics are presented in Table 1, Panel A. As shown in Figures 1 and 2, when readability is high, there is no discernable effect of language sentiment (essentially a flat line in both

---

[5] Neither the main effect of readability nor the interaction of readability and language sentiment is significant for the language sentiment manipulation check questions (smallest $p = 0.12$). We also analyze the three language sentiment manipulation check questions for the less and more sophisticated investor groups, respectively, and find similar results ($p < 0.01$ for all three measures in both investor groups).

[6] Neither the main effect of language sentiment nor the interaction of readability and language sentiment is significant for the manipulation check questions on readability (smallest $p = 0.85$). We also conduct separate readability manipulation checks for both the less and more sophisticated investor groups, and find similar results ($p = 0.06/0.02/0.57$ for the difficult to read/understand/process measures in the less-sophisticated investor group; $p < 0.01/p < 0.01/p = 0.05$ for the difficult to read/understand/process measures in the more-sophisticated investor group).

figures) for both less sophisticated investors and more sophisticated investors. In contrast, when readability is low, the directional impact of language sentiment varies with investor sophistication level. Specifically, language sentiment has a positive slope for less sophisticated investors but a negative slope for more sophisticated investors.[7]

To test our hypotheses, we conduct a three-way ANOVA analysis with investor sophistication level, language sentiment, and disclosure readability as the independent variables, and investors' assessments of the firm's future earnings performance as the dependent variable. The ANOVA results are shown in Table 1, Panel B. We find that, overall, investors' earnings judgments are significantly higher when readability is low (0.65) than when it is high (-0.06), $p = 0.05$. Insofar as the firm's trend performance described in the key financial highlights section (where the readability manipulation occurs) is disappointing, this main effect of readability suggests that poor earnings performance is more likely to be obfuscated when the earnings release is difficult to read than when it is easy to read. The obfuscating role of low readability is more salient when investors are less sophisticated than more sophisticated, as indicated by the significant interaction effect of readability and investor sophistication ($p = 0.05$). Specifically, compared to high readability, low readability leads to significantly higher earnings judgments for less sophisticated investors (-0.41 vs. 1.00, $p < 0.01$), but not for more sophisticated investors (0.27 vs. 0.26, $p = 0.99$). We also find a significant interaction effect of language sentiment and investor sophistication ($p = 0.01$), in that positive language leads to higher earnings judgments for less sophisticated investors (0.77 vs. -0.11, $p = 0.05$, one-tailed), but lower earnings judgments for more

---

[7] For each experimental condition, we also classify participants into two groups: those who believe that earnings performance will be strong in the future (i.e., responses ranging from "+1" to "+5") and those who believe otherwise (i.e., responses ranging from "-5" to "0"). We then calculate the number of participants in the "strong" category as a percentage of the total number of participants in that condition (hereafter, %Strong). We obtain similar results to our main analysis. With respect to less sophisticated investors, we find that %Strong is significantly higher in the positive language condition than that in the neutral language condition when readability is low (64.3% vs. 35.7%, $p < 0.01$), but not significantly different from each other when readability is high (53.3% vs. 46.7%, $p = 0.52$). With respect to more sophisticated investors, we find that %Strong is significantly lower in the positive language condition than that in the neutral language condition (38.1% vs. 61.9%, $p = 0.08$), but not significantly different from each other when readability is high (31.6% vs. 44.4%, $p = 0.42$).

sophisticated investors (-0.25 vs. 0.80, p = 0.02, one-tailed).[8] In addition, our results also show a significant three-way interaction effect (p = 0.01). We next examine how readability and language sentiment influence the judgments of less sophisticated investors and more sophisticated investors, respectively.

[Insert Table 1, Figure 1, and Figure 2 about here]

4.2.2 Less sophisticated Investors

H1 posits that positive language will result in higher earnings judgments for less sophisticated investors, and that this effect will be more pronounced when readability is low. The results in Table 2 and Figure 1 are generally consistent with H1. We find that less sophisticated participants rate future earnings more favorably when the earnings release is difficult to read and positively framed (See Panel A of Table 2). Specifically, the average earnings judgments in the low readability/positive language condition is 1.74 (significantly greater than 0, p < 0.01, indicating a belief in strong future performance), while the average evaluations in the high readability/positive language, high readability/neutral language, and low readability/neutral language condition are -0.38, -0.44, and 0.22, respectively (not significantly different from 0, smallest p = 0.39, indicating a neutral opinion on future performance), respectively. These results suggest that only participants in the low readability/positive language condition believe that W&Z's earnings performance will be strong in the future.

[Insert Table 2 about here]

To test whether less sophisticated participants' earnings judgments vary in accordance with H1, we first conduct an analysis of variance (ANOVA) with language sentiment and readability as the independent variables, and earnings judgments as the dependent variable (see Panel B of Table 2). We observe a significant main effect of readability (-0.41 vs. 1.00, p < 0.01), a significant main effect of language sentiment

---

[8] To estimate the effect size of the joint effect of language sentiment/readability, we compute the percentage increase in earnings judgments between key experimental conditions. Specifically, we first convert our response scale of "-5" to "+5" to a scale of "0" to "10." Our numerator is the difference in earnings judgments between two conditions, and the denominator is the smaller of the two earnings judgments. For example, for less sophisticated investors, when readability is low, the earnings judgments in the positive language condition are 29.11% higher than those in the neutral language condition. In contrast, for more sophisticated investors, when readability is low, the earnings judgments in the neutral language condition are 55.58% higher than those in the positive language condition.

(0.77 vs. -0.11, p = 0.05, one-tailed), and a marginally significant interaction effect (p = 0.07, one-tailed). As H1 implies an ordinal interaction, a contrast test is more powerful (Buckless and Ravenscroft [1990], Rosnow and Rosenthal [1995]). Accordingly, we also conduct a contrast test with the following contrast weights: +3 in the Positive Language/Low Readability Condition, +1 in the Neutral Language/Low Readability Condition, -2 in the Positive Language/High Readability Condition, and -2 in the Neutral Language/High Readability Condition. These contrast weights reflect our prediction that positive language leads to more favorable earnings judgments, and that its effect is more pronounced when readability is low. The results for the contrast test is significant (p = 0.01), as shown in Panel C of Table 2. Thus, H1 is supported.

Follow-up tests (see Panel D of Table 2) show that when readability is low, participants' earnings judgments are significantly higher in the positive versus neutral language sentiment condition (1.74 vs. 0.22, p = 0.03). However, when readability is high, participants' earnings judgments are not significantly affected by language sentiment (-0.38 vs. -0.44, p = 0.92). Taken together, these results provide support for our theory that the use of positive language to convey a favorable earnings prospect to investors will be effective for less sophisticated investors only when the message is difficult to read. In addition, we find that low readability influences less sophisticated investors' judgments only when the language is positive (-0.38 vs. 1.74, p < 0.01), but not when the language is neutral (-0.44 vs. 0.22, p = 0.32). These results suggest that the effects of readability on less sophisticated investors' earnings judgments are also contingent on language sentiment.

4.2.3 More Sophisticated Investors

H2 predicts that more sophisticated investors' evaluations of firm's future earnings performance are more favorable when earnings releases are neutrally written than when they are positively written, and that the difference is larger when earnings releases are difficult to read. We first conduct an analysis of variance (ANOVA) with language sentiment and readability as independent variables, and earnings judgments

as the dependent variable (see Panel B of Table 3). We observe a significant main effect of language sentiment (-0.25 vs. 0.80, p = 0.02, one-tailed), an insignificant main effect of readability (0.27 vs. 0.26, p = 0.99), and a significant interaction effect of language sentiment and readability (p = 0.01, one-tailed). Follow-up tests (see Panel C of Table 3) show that when readability is low, earnings judgments are higher if the language is neutral than when it is positive (1.41 vs. -0.88, p < 0.01). In contrast, when readability is high, there is no difference in participants' earnings judgments between neutral and positive language conditions (0.22 vs. 0.32, p = 0.89). These results support our argument that more sophisticated investors will lower their earnings judgments if the language is positive, especially when readability is low.

In addition, when the language is positive, participants in the high readability condition make marginally higher earnings judgments than those in the low readability condition (0.32 vs. -0.88, p = 0.10). In contrast, when the language is neutral, participants in the high readability condition make marginally lower earnings judgments than those in the low readability condition (0.22 vs. 1.41, p = 0.10). These results suggest that the directional effects of readability on more sophisticated investors' earnings judgments are conditional on language sentiment.

[Insert Table 3 about here]

We also compare the earnings judgments between more and less sophisticated investors within each condition. We find that when readability is high, there is no significant difference in earnings judgments between more and less sophisticated investors, either in the positive language condition (0.32 vs. -0.38, p = 0.33) or in the neutral language condition (0.22 vs. -0.44, p = 0.34). However, this situation changes when readability is low. Compared to less sophisticated investors, more sophisticated investors' earnings judgments are significantly lower in the positive language condition (-0.88 vs. 1.74, p < 0.01, one-tailed) but higher in the neutral language condition (1.41 vs. 0.22, p = 0.05, one-tailed). In fact, the difference in earnings judgments between more and less sophisticated investors in the low readability/positive language condition (-2.62) is significantly larger than that in the

low readability/neutral language condition (1.19), $p < 0.01$. These results suggest that low readability disclosures increase information asymmetry between investors who vary in sophistication, an effect that is magnified with disclosures couched in positive language.

4.3 Additional Analysis

4.3.1 The Attribute Framing Effects of Positive Language

To investigate whether positive language leads to positive framing effects, we explore how language sentiment affects participants' initial impressions of management's comments on the firm's performance. These management comments are discussed in the first two paragraphs of the earnings release. We ask participants to rate the favorableness of four performance indicators compared to previously issued management guidance (e.g., net sales, earnings per share, earnings per share from continuing operations, and overall firm performance) on a 5-point scale (1-Poor, 2-Neutral, 3-Good, 4-Very Good, 5-Outstanding). Across both less sophisticated and more sophisticated investors, we find that, compared to neutral language, positive language leads to more favorable assessments of the same numerical indicators for net sales ($p = 0.06/0.08$ for less/more sophisticated investors), earnings per share ($p < 0.01/p = 0.08$ for less/more sophisticated investors), earnings per share from continuing operations ($p < 0.01$ for both less and more sophisticated investors), and overall firm performance ($p = 0.01$ for both less and more sophisticated investors).[9] In addition, we ask participants to describe management's confidence in the company's future performance on an 11-point Likert scale, with "0" indicating "Not at all Confident" and "10" indicating "Very Confident." We find a significant main effect of language sentiment ($p = 0.08/0.06$, one-tailed, for less/more sophisticated investors), indicating that participants perceive positive language to reflect more management confidence in the future. For all variables in this analysis, we find no main effect of readability or investor sophistication ($p > 0.26$), or other interaction

---

[9] We also ask participants to describe the external economic conditions and W&Z's operating cash flow. The effect of positive language is insignificant for both less and more sophisticated participants ($p > 0.31$).

effects (p > 0.17). Overall, our results suggest that positive as opposed to neutral language leads investors to perceive the firm's performance (relative to prior management guidance) to be better, and management's confidence to be higher, irrespective of readability and investor sophistication.

4.3.2 The Role of Understanding on Less Sophisticated Investors' Judgments

Inconsistent with the positive language sentiment in the first two paragraphs of the earnings release, the majority of the performance indicators discussed in the key financial highlights section are worse than those in the year-ago quarter (see Appendix 2). For example, net sales dropped 6%, earnings per share dropped 1%, and earnings from continuing operations dropped 3%. We evaluate whether participants understand that the firm's performance compared with year-ago quarter is unfavorable (hereafter, understanding) by assessing their responses to the above three items. Correct answers are coded into "1", and incorrect answers are coded into "0." Adding up participants' responses to changes of net sales, earnings per share, and earnings from continuing operations, our understanding measure ranges from "0" (all wrong responses) to "3" (all correct responses).

To examine whether positive language influences less sophisticated investors' understanding, we conduct an ANOVA analysis with readability and language sentiment as independent variables, and understanding as the dependent variable. For less sophisticated investors, we find that the main effect of language sentiment is marginally significant (means = 1.43 and 1.86 for positive and neutral language, respectively; p = 0.06, one-tailed). These results suggest that positive language is more likely to impair less sophisticated investors' understanding of the unfavorable trend performance than neutral language. Neither the main effect of readability nor the interaction effect is significant (p > 0.73).[10]

---

[10] We also ask participants to assess changes of earnings from discontinued operations, guidance on future core sales growth, and guidance on future earnings per share, all of which reflect positive changes. We code a separate measure for participants' understanding of these positive changes (i.e., understanding-positive) and a separate measure for participants' understanding of both negative and positive changes (understanding-total). Results are similar if we use the "understanding-positive" measure (means = 1.31 and 1.75 for positive and neutral language, respectively; p = 0.03, one-tailed) or the "understanding-total" measure (means = 2.74 and 3.61 for positive and neutral language, respectively; p = 0.03, one-tailed).

To further investigate whether less sophisticated investors' understanding mediates their judgments on the firm's future earnings performance, we conduct a structural equation modeling analysis with language sentiment as the independent variable, readability as the moderator, understanding as the intervening variable, and investors' earnings judgments as the dependent variable. The overall model describes the relationships in the data well.[11] As shown in Figure 3, we find that positive language decreases investors' understanding (coefficient = -0.19, p = 0.05, one-tailed), and that this reduced understanding of the unfavorable trend performance leads to more optimistic earnings judgments (coefficient = -0.23, p < 0.01). There is no interaction effect of readability and language sentiment on understanding ($\chi^2 = 0.02$, p = 0.88). These results provide further evidence that language sentiment "switches off" less sophisticated investors' analytical mode (i.e., understanding), subjecting such investors to the positive heuristic sentiment effect of positive language.

[Insert Figure 3 about here]

4.3.2 Credibility Influences on More Sophisticated Investors' Judgments

With respect to more sophisticated investors, our results show that their understanding is neither influenced by the positive framing effects arising from positive language (p = 0.31 for the main effect of language sentiment) nor the obfuscation effect arising from low readability (p = 0.47 for the main effect of readability).[12] Hence, neither positive language sentiment nor low readability turns off sophisticated investors' analytical mode (in terms of understanding). We then investigate why more sophisticated investors estimate *lower* future performance when they are reading earnings release with positive language. We posit that when management's self-comments on the firm's performance are inconsistent with the key financial highlights, variations in readability and language sentiment are more likely to cause more sophisticated investors to assess the credibility of the earnings release,

---

[11] We confirm the model's goodness of fit with a conventional χ^2 test ($\chi^2 = 8.18$, p = 0.32), an Incremental Fit Index of 0.92 (which is above the cutoff of 0.90 for model acceptance [Byrne 2001]), and a Root Mean Square Error of Approximation of 0.04 (which is below the cut-off point of 0.08 for a good fit [MacCallum, Browne, and Sugawara 1996]).
[12] Results are similar if we use the "understanding-positive" measure (p > 0.52) or the "understanding-total" measure (p > 0.26).

as those linguistic features may implicitly reflect management's strategic intentions.

In order to validate this argument, we ask participants to evaluate the credibility and believability of the earnings release on an 11-point Likert scale ranging from 0 (Not at all Credible/Believable) to 10 (Very Credible/Believable). The two measures are highly correlated (Pearson correlation = 0.93, $p < 0.01$; Cronbach's alpha = 0.96), so we use the average of these two measures as our measure of earnings release credibility. We first conduct ANOVA analyses with language sentiment and readability as the independent variables, and earnings release credibility as the dependent variable for more sophisticated investors. We find a significant main effect of language sentiment ($p < 0.01$) and a significant interaction effect ($p = 0.02$). Specifically, positive language leads to lower credibility assessment when readability is low (4.82 vs. 7.24, $p < 0.01$), but not when readability is high (5.92 vs. 6.25, $p = 0.57$). These results support our argument that more sophisticated investors tend to find low-readability earnings releases couched in positive language to be less credible. The findings also demonstrate that more sophisticated investors employ greater in-depth information processing, as they are able to infer management's strategic intentions.[13]

Next, we investigate whether earnings release credibility mediates the effect of language sentiment and readability on more sophisticated investors' earnings judgments. As shown in Figure 4, we conduct a structural equation modeling analysis with language sentiment as the independent variable, readability as the moderator, earnings release credibility as the intervening variable, and investors' earnings judgments as the dependent variable. The overall model describes the relationships in

---

[13] Compared with more sophisticated investors, less sophisticated investors are less likely to take the credibility of the earnings release into consideration. Consistent with this argument, we find that for less sophisticated investors' credibility assessments, the main effects of readability and language sentiment along with the interaction effect are all insignificant ($p > 0.40$). In addition, we also ask participants to indicate the extent that they think the content of the earnings release is misleading on an 11-point scale ranging from 0 (not at all misleading) to 10 (very misleading). We find that when readability is low, more sophisticated investors consider disclosures with neutral language to be less misleading than those with positive language (4.53 vs. 6.06, $p = 0.05$), but no such language sentiment effect is found for less sophisticated investors (5.26 vs. 5.33, $p = 0.92$). This provides additional evidence that more sophisticated investors are more likely to engage in analytical processing than less sophisticated investors.

the data well.[14] We find that when readability is low, positive language results in lower earnings release credibility (coefficient = -0.56, p < 0.01). In contrast, when readability is high, positive language does not influence more sophisticated investors' assessment of earnings release credibility (coefficient = -0.01, p = 0.56). The interaction effect of readability and language sentiment on earnings release credibility is significant (χ^2 = 6.00, p = 0.01). Moreover, earnings release credibility positively influences more sophisticated investors' earnings judgments (coefficient = 0.33, p < 0.01). These results show that earnings release credibility mediates the joint effect of language sentiment and readability on more sophisticated investors' earnings judgments.[15]

[Insert Figure 4 about here]

5. Discussion and Conclusion

We investigate how the effects of language sentiment on investor judgment are moderated by readability and investor sophistication level. In an experiment with MBA students as participants, we show that when readability is high, language sentiment does not affect investors' earnings judgments, irrespective of investor sophistication level. When readability is low, less sophisticated investors make higher earnings judgments when they are given an earnings release with positive language. Additional analysis shows that positive language is more likely to impair less sophisticated investors' understanding of firm's unfavorable trend performance. As a result, less sophisticated investors tend to make judgments based on the favorable

---

[14] We confirm the model's goodness of fit with a conventional χ^2 test (χ^2 = 5.79, p = 0.33), an Incremental Fit Index of 0.98, which is above the cutoff of 0.95 for a well-fit model (Byrne [2001]), and a Root Mean Square Error of Approximation of 0.03, which is below the cut-off point of 0.08 for a good (MacCallum, Browne, and Sugawara [1996]).

[15] As shown in Tables 1 and 3, among more sophisticated investors, with neutral language, earnings judgments are marginally higher in the low readability condition than in the high readability condition (p = 0.05, one-tailed); the opposite effect is obtained with positive language (p = 0.05, one-tailed). These results could be explained by the moderated mediating role of credibility assessment on more sophisticated investors' earnings judgments. As stated above, language sentiment influences credibility assessment only when readability is low. Hence, participants in the low (but not high) readability condition will accord a credibility premium associated with neutral as opposed to positive language. This difference in credibility premium in turn results in higher earnings judgments in the low (compared to high) readability condition when language is neutral. Similarly, when language is positive, participants in the low readability condition are more likely to penalize positive language, which leads to lower earnings judgments compared to those in the high readability condition.

impression induced by positive language, which causes them to perceive the firm's earnings performance to be stronger in the future. In contrast, more sophisticated investors make lower earnings judgments when readability is low and the earnings release is written in positive language. Further analysis demonstrates that this effect is mediated by more sophisticated investors' concerns for the credibility of the earnings release. Specifically, when more sophisticated investors find that the positive language in the earnings release is not supported by some aspects of the disclosed performance information, they perceive the earnings release to be less credible. As a result, their assessments of future earnings performance drop because they find the earnings release to be less credible.

Our study contributes to recent literature investigating the impact of disclosure/management tone on investors' reactions. While existing studies cannot rule out the argument that investors' reactions toward tone are partially due to its non-informational role, we provide evidence of a language sentiment effect in that the use of positive language can significantly influence investors' perceptions and evaluations of financial performance, even when hard quantitative information is held constant. More importantly, rather than purely investigating the impact of language sentiment, we take both the target feature (readability of disclosures) and the judge feature (sophistication level of investors) into consideration. As a departure from previous studies that explore the general main effect of tone, we show that the impact of language sentiment is contingent on both disclosure readability and investor sophistication. Thus, we provide a more complete picture regarding how language sentiment plays a role in investors' judgments.

Interestingly, while we find that positive language leads to positive earnings judgments from less sophisticated investors but negative earnings judgments from more sophisticated investors, the overall market reaction to positive tone is positive, as documented in the current archival literature (Davis et al. [2012], Davis, Piger, and Sedor [2012], Davis and Tama-Sweet [2012]). One interpretation, therefore, is that prior archival studies could have had difficulties in isolating the effect of pure positive

language from the positive information content in the disclosures. Future research can examine how market reactions of less and more sophisticated investors aggregate to form the overall market reaction, and the conditions under which less (or more) sophisticated investors' reactions are more influential in the overall market. A market experiment could be helpful for such an investigation (e.g., see Libby, Bloomfield, and Nelson [2002]).

Other than the literature on language sentiment, our study also contributes to the stream of literature on disclosure readability by demonstrating that its effect on investors' earnings judgments is contingent on language sentiment and investor sophistication. Neither moderator has been examined jointly or in isolation in prior studies on readability effects. Our results suggest that deleterious effects of low readability occur with positive tone among less sophisticated investors. Importantly, we show that the deleterious effects of low readability actually reverse with more sophisticated investors when the disclosure is written with positive language. Thus, more sophisticated investors are less likely to be obfuscated by low readability and they can undo the framing effects arising from positive language. This finding is particularly striking since our more sophisticated investor participants are actually MBA students, not financial analysts who are investors with higher sophistication. This suggests that an important and non-trivial segment of the general non-professional investor population is able to see through and undo the effects of low readability and positive language sentiment. Consequently, it may be that these adverse effects of low readability documented in prior research (e.g., Miller [2010], Rennekamp [2012], You and Zhang [2009]) do not generalize even among this general non-professional group of investors.

Our study only examines the effect of language sentiment and readability in a context where management provides an inconsistent message in an earnings release. Specifically, while the reported EPS beats management earnings guidance, it is worse than that in the year-ago quarter. We choose such a context because management is more likely to use positive language and low readability to obfuscate investors in the

face of unsatisfactory earnings performance (Bloomfield [2002]). Although we expect that the documented role of readability will be similar but less pronounced when message is consistent, it is not clear how message consistency will moderate the main effect of language sentiment and its interaction with investor sophistication level. Future research can investigate the impact of language sentiment and investor sophistication level when the underlying message is consistently positive or negative. In addition, rather than manipulate both readability and language sentiment in all four paragraphs of the earnings release, we manipulate language sentiment in the first two paragraphs and readability in the last two paragraphs. We believe that our current design enables us to have clean manipulations of language sentiment and readability. However, we cannot address whether the results would change if readability and language sentiment were manipulated for the same earnings release content. The challenge in such a manipulation is that readability can influence the salience of the language sentiment and vice versa, precluding both independent variables from truly being orthogonal. Finally, although we proxy investor sophistication in terms of whether MBA students have more accounting and finance knowledge/frequency of reading annual reports and earnings releases, we acknowledge that MBA students are not necessarily representative of the lower/upper ends of this continuum of investor sophistication. For instance, less sophisticated investors may have much less financial literacy, while more sophisticated investors may be better represented by financial analysts.

REFERENCES

BLOOMFILED, R. "The 'Incomplete Revelation Hypothesis' and Financial Reporting." *Accounting Horizons* 16 (2002): 233-243.

BLOOMFIELD, R., AND P. E. FISCHER. "Disagreement and the Cost of Capital." *Journal of Accounting Research* 49 (2011): 41 – 68.

BLOOMFIELD, R.; R. LIBBY; AND M. W. NELSON. "Communication of Confidence as a Determinant of Group Judgment Accuracy." *Organizational Behavior and Human Decision Process* 68 (1996): 287-300.

BUCKLESS, F. A.. AND S. P. RAVENSCROFT. "Contrast coding: A Refinement of ANOVA in Behavioral Analysis." *The Accounting Review* 64 (1990): 933-945.

BYRNE, B. M. Structural Equation Modeling With AMOS: Basic Concepts, Applications, and Programming. Mahwah, NJ: Lawrence Erlbaum Associates. 2001.

CARMINES, E., AND R. ZELLER. Reliability and Validity Assessment. Sage Publications, Beverly Hills, California. 1979.

CHAIKEN, S. "Heuristic Versus Systematic Information Processing And The Use Of Source Versus Message Cues In Persuasion." *Journal of Personality and Social Psychology* 39 (1980): 752-766.

CHAIKEN, S.; A. LIBERMAN; AND A. H. EAGLY. "Heuristic And Systematic Processing Within And Beyond The Persuasion Context." In J. S. Uleman and J.A. Bargh (Eds.), Unintended thought (pp.212-252). New York: Guilford. 1989.

CUTTS, M. The Plain English Guide. Oxford University Press. 1996.

DAVIS, A; W. L. GE; D. MATSUMOTO; AND J. L. ZHANG. "The Effect of Managerial "Style" on the Tone of Earnings Conference Calls." *Working Paper.* 2012

DAVIS, A.; J. PIGER; AND L. SEDOR. "Beyond the Numbers: Measuring the Information Content of Earnings Press Release Language." *Contemporary Accounting Research* 29 (2012): 845-868.

DAVIS, A., AND I. TAMA-SWEET. "Managers' Use of Language Across Alternative Disclosure Outlets: Earnings Press Releases Versus MD&A." *Contemporary Accounting Research* 29 (2012): 804-837.

ENGELBERG, J. "Costly Information Processing: Evidence from Earnings Announcement." *Working paper*, University of California, San Diego, 2008

ELLIOTT, W. B.; F. HODGE; J. KENNEDY; AND M. PRONK. "Are M.B. A. Students a Good Proxy for Nonprofessional Investors?" *The Accounting Review* 82 (2007): 139-168.

FAIRFIELD, P.; R. J. SWEENEY; AND T. L. YOHN. "Accounting Classification and the Predictive Content of Earnings." *The Accounting Review* 71 (1996): 337 – 55.

FELDMAN, R.; S. GOVINDARAJ; J. LIVNAT; AND B. SEGAL. "The Incremental Information Content of Tone Change in Management Discussion and Analysis." *Working Paper*, 2009

FELDMAN, R.; S. GOVINDARAJ; J. LIVNAT; AND B. SEGAL. "Management's Tone Change, Post Earnings Announcement Drift and Accruals." *Review of*

Accounting Studies 15 (2010): 915–53.

GORN, G. J. "The Effects of Music in Advertising on Choice Behavior: A Classical Conditioning Approach." *Journal of Marketing* 46 (1982): 94-101.

HALES, J.; J. KUANG; AND S. VENKATARAMAN. "Who Believes the Hype? An Experimental Examination of How Language Affects Investor Judgments." *Journal of Accounting Research* 49 (2011): 223–55.

HAN, J., AND H. T. TAN. "Investors' Reactions to Management Earnings Guidance: The Joint Effect of Investment Position, News Valence, and Guidance Form." *Journal of Accounting Research* 48 (2010): 81-104.

HENRY, E. "Are Investors Influenced by the Way Earnings Press Releases are Written?" *Journal of Business Communication* 45 (2008): 363-408.

HOSSAIN, T., AND J. A. LIST. "The Behavioralist Visits the Factory: Increasing Productivity Using Simple Framing Manipulations." *Management Science* 58 (2012): 2151-2167.

HUANG, X.; S. H. TEOH; AND Y. L. ZHANG. "Tone management." *-The Accounting Review* (2013), forthcoming.

HUTTON, A. P.; G. S. MILLER; AND D. J. SKINNER. "The role of supplementary statements with management earnings forecasts." *Journal of Accounting Research* 41 (2003): 867-90.

KUVAAS, B., AND M. SELART. "Effects of Attribute Framing on Cognitive Processing on Evaluation." *Organizational Behavior and Human Decision Processes* 95 (2004): 198-207.

LARCKER, D. F., AND B. TAYAN. "Financial Manipulation: Words Don't Lie" Rock Center for Corporate Governance at Stanford University Closer Look Series: Topics, Issues and Controversies in Corporate Governance No. CGRP-07, 2010.

LEVIN, I. P., AND G. J. GAETH. "How Consumers Are Affected By The Framing Of Attribute Information Before And After Consuming The Product." *Journal of Consumer Research* 15 (1988): 374-78.

LEVIN, I. P.; S. L. SCHNEIDER; AND G. J. GAETH. "All Frames are not Created Equal: A Typology and Critical Analysis of Framing Effects." *Organizational Behavior and Human Decision Processes* 76 (1998): 149-88.

LI, F. "Annual Report Readability, Current Earnings, and Earnings Persistence." *Journal of Accounting and Economics* 45 (2008): 221–47.

LI, F. "The Information Content of Forward-Looking Statements in Corporate Filings—A Naïve Bayesian Machine Learning Approach." *Journal of Accounting Research* 48 (2010):1049–1102.

LIBBY, R.; R. BLOOMFIELD; AND M. W. NELSON. "Experimental Research in Financial Accounting." *Accounting, Organizations and Society* 23 (2002): 775-810.

LOUGHRAN, T., AND B. MCDONALD. "Measuring Readability in Financial Text." *Working paper*, The University of Notre Dame. 2010.

MACCALLUM, R. C.; M. W. BROWNE; AND H. M. SUGAWARA. "Power Analysis and Determination of Sample Size for Covariance Structure Modeling" *Psychological Method* 1 (1996): 130-149.

MACGREGOR, D. G.; P. SLOVIC; D. DREMAN; AND M. BERRY. "Imagery, Affect, and Financial Judgment." *Journal of Psychology and Financial Markets* 1 (2000): 104-110.

MCELROY, T., AND J. J. SETA. "Framing Effects: An Analytic-Holistic Perspective." *Journal of Experimental Social Psychology* 39 (2003): 610-17.

MILLER, B. "The Effects of Reporting Complexity on Small and Large Investor Trading." *The Accounting Review* 85 (2010): 2107-2143.

PETTY, R. E., AND J. T. CACIOPPO. "Communication And Persuasion: Central And Peripheral Routes To Attitude Change." New York: Springer-Verlag. 1986.

PETTY, R. E.; D. W. SCHUMANN; S. A. RICHMAN; AND A. J. STRATHMAN. "Positive Mood And Persuasion: Different Roles For Affect Under High And Low Elaboration Conditions." *Journal of Personality and Social Psychology* 64(1993): 5–20.

RENNEKAMP, K. "Processing Fluency and Investors' Reactions to Disclosure Readability." *Journal of Accounting Research* 50 (2012): 1319-1354.

ROGERS, J. L.; A. VAN BUSKIRK; AND S. L. C. ZECHMAN. "Disclosure Tone and Shareholder Litigation" *The Accounting Review* 86 (2011): 2155-2183.

ROSNOW, R. L., AND R. ROSENTHAL. "Some Things You Learn Aren't So: Cohen's Paradox, Asch's Paradigm, and the Interpretation of Interaction." Psychological Science 6 (1995): 3-9.

SECURITIES AND EXCHANGE COMMISSION. "A plain English Handbook: How to create clear SEC disclosure." SEC Office of Investor Education and Assistance. 1998. Retrieved from the SEC's website: http://www.sec.gov/pdf/handbook.pdf

TAN, H. T.; E. Y. WANG; AND B. ZHOU. "How Does Readability Influence Investors' Judgments? Consistency of Benchmark Performance Matters." *Working Paper*, 2012.

TETLOCK, P. C. "Giving Content to Investor Sentiment: The Role of Media in the Stock Market." *Journal of Finance* 62 (2007): 1139-1168.

TVERSKY, A. AND D. KAHNEMAN. "Rational Choice and the Framing of Decisions." *The Journal of Business* 59 (1986): S251-S278.

YOU, H., AND X. ZHANG. "Financial Reporting Complexity and Investor under-reaction to 10-K Information." *Review of Accounting Studies* 14 (2009): 559-586.

WEGENER, D., AND R. PETTY. "Flexible Correction Processes in Social Judgment: The Role of Naïve Theories in Corrections for Perceived Bias." *Journal of Personality and Social Psychology* 68 (1995): 36-51.

APPENDIX 1

Manipulation of Language Sentiment

POSITIVE treatment

W&Z reported strong second quarter results that achieved the company's overall expectations. Net sales were $19.8 billion, far above company's guidance of $18.9 billion. Earnings per share were $1.08, significantly higher than the low end of company's guidance range of $1.02 to $1.09. Earnings per share from continuing operations were $1.00 versus prior guidance of $1.01. To reflect the greater growth projection in earnings per share, the company favorably increased the low end of previous guidance range by adding $0.03 per share.

"We are very pleased with our second quarter results," said President Bill Sloan, "We made the right choice to focus on the fundamentals that are critical to success in our business. Despite the fact that the economic condition remained somewhat weak and volatile, I am very impressed by our core sales growth performance of up two percent, which is sharply higher than our prior guidance of minus three percent. I am also very pleased to see that earnings per share reached $1.08, which were almost the highest end of our expectations. Earnings per share from continuing operations were one cent less than previous guidance. Our cash position remained stable in the first and second quarter. Operating cash flow was $2.3 billion for the first quarter and $2.2 billion for the second quarter. Encouraged by our strong performance in the second quarter, we have positively raised core sales growth guidance by one percent for the coming quarter."

NEUTRAL treatment

W&Z reported second quarter results that were roughly in line with the company's overall expectations. Net sales were $19.8 billion, relative to company's guidance of $18.9 billion. Earnings per share were $1.08, within the company's guidance range of $1.02 to $1.09. Earnings per share from continuing operations were $1.00, lower than prior guidance of $1.01. To reflect the revised growth projection in earnings per share, the company updated the low end of previous guidance range by adding $0.03 per share.

"The second quarter results met our overall expectations," said President Bill Sloan, "We focused on the fundamentals that are critical to our business. The economic condition remained somewhat weak and volatile. Core sales growth was up two percent versus our prior guidance of minus three percent. Earnings per share of $1.08 were in line with our expectations. However, I am disappointed that earnings per share from continuing operations failed to meet our previous guidance by one cent. Though our cash position remained stable in the first and second quarter, operating cash flow dropped from $2.3 billion for the first quarter to $2.2 billion for the second quarter. Based on our performance in the second quarter, we have adjusted core sales growth guidance by adding one percent for the coming quarter."

APPENDIX 2

Manipulation of Readability

<u>HIGH READABILITY</u> treatment

Key Financial Highlights

<table>
<tr><th colspan="4">Quarterly Results Summary</th></tr>
<tr><th></th><th>*2010 Q2*</th><th>*VS. 2009 Q2*</th><th>*Sources/Reasons*</th></tr>
<tr><td>Net sales</td><td>$19.8 billion</td><td>down 6%</td><td>Foreign exchange impacts of 7% as U.S. dollar strengthened</td></tr>
<tr><td>Unit sales volume</td><td>N/A</td><td>down 3%</td><td>General economic conditions, market contractions, prior year divestitures and share losses</td></tr>
<tr><td>Gross margin</td><td>52.6%</td><td>up 2.9%</td><td>Price increases, lower commodity costs and manufacturing cost savings</td></tr>
<tr><td>Operating margin</td><td>22.5%</td><td>down 1.6%</td><td>Higher SG&A expenses as a percentage of net sales, partially offset by higher gross margin</td></tr>
<tr><td>Earnings per share</td><td>$1.08</td><td>down 1%</td><td>Lower net sales, partially offset by the gain on sale of Acxon in Japan</td></tr>
<tr><td>Earnings from discontinued operations</td><td>$0.24 billion</td><td>up 2%</td><td>Gain on sale of Acxon in Japan</td></tr>
<tr><td>Earnings from continuing operations</td><td>$3.02 billion</td><td>down 3%</td><td>Foreign exchange impacts as U.S. dollar strengthened, lower net sales, and higher base-period divestiture gains</td></tr>
</table>

2010 Third Quarter Guidance

- Core sales growth: 2-4%, compared with previous forecast of 1-3%
- Net sales: up 3-6% versus year-ago quarter
- Foreign exchange impact: contributes 1-2% to net sales growth
- Earnings per share: $1.25-$1.35, compared with previous forecast of $1.22-$1.35

Key Financial Highlights

Compared with year-ago quarter, net sales declined six percent to $19.8 billion for the April-June quarter mainly due to foreign exchange impacts of seven percent as the U.S. dollar strengthened. Unit sales volume decreased three percent largely due to general economic conditions, market contractions, prior year divestitures and share losses. Gross margin expanded 2.9 percent to 52.6 percent behind price increases, lower commodity costs and manufacturing cost savings. Operating margin was 22.5 percent, a decrease of 1.6 percent versus the prior year period driven by higher SG&A expenses as a percentage of net sales, partially offset by higher gross margin. Earnings per share were $1.08, a decrease of one percent primarily due to lower net sales, partially offset by the gain on the sale of Acxon in Japan, which increased earnings from discontinued operations by two percent to $0.24 billion. Earnings from continuing operations were $3.02 billion, down three percent behind foreign exchange impacts as the U.S. dollar strengthened, lower net sales, and higher base-period divestiture gains.

2010 Third Quarter Guidance

For the third quarter, the company increased the range of expected core sales growth by one percent to plus two to four percent. Net sales are expected to be up three to six percent versus year-ago quarter. Foreign exchange is expected to contribute one to two percent to net sales growth. The company updated its earnings per share guidance to $1.25 to $1.35 by increasing the low end of the previous guidance ranges by $0.03/share.

TABLE 1

All Participants' Earnings Judgments

DV = Future Earnings Performance Prediction

To what extent do you agree that W&Z's **earnings performance will be strong** in the near future?

| -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 |

Strongly DISAGREE   Neutral   Strongly AGREE

Panel A: Descriptive Statistics – Mean (Standard Deviation) [Sample Size]

Language Sentiment	Readability				
	High		*Low*		
	Neutral	Positive	Neutral	Positive	Overall
Less Sophisticated	-0.44 (2.12) [18]	-0.38 (2.45) [16]	0.22 (1.80) [18]	1.74 (1.63) [19]	0.27 (1.98) [71]
More Sophisticated	0.22 (2.29) [18]	0.32 (1.70) [19]	1.41 (2.35) [17]	-0.88 (2.15) [17]	0.26 (2.23) [71]
Overall	-0.06 (2.12) [71]		0.65 (2.20) [71]		

Panel B: Three-Way ANOVA Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Sophistication	0.01	1	0.01	0.00	0.96
Readability	16.97	1	16.97	3.98	0.05
Language Sentiment	0.84	1	0.84	0.20	0.66
Sophistication * Readability	17.18	1	17.18	4.03	0.05
Sophistication * Language Sentiment	31.68	1	31.68	7.42	0.01
Readability * Language Sentiment	1.97	1	1.97	0.46	0.50
Sophistication*Readability*Language Sentiment	32.49	1	32.49	7.61	0.01
Error	572.09	134	4.27		

TABLE 2

Less Sophisticated Participants' Earnings Judgments

DV = Future Earnings Performance Prediction

To what extent do you agree that W&Z's **earnings performance will be strong** in the near future?

-5	-4	-3	-2	-1	0	1	2	3	4	5

Strongly DISAGREE   Neutral   Strongly AGREE

Panel A: Descriptive Statistics – Mean (Standard Deviation) [Sample Size]

Language	Readability								
Sentiment		High			Low			Overall	
Positive	-0.38	(2.45)	[16]	1.74	(1.63)	[19]	0.77	(2.28)	[35]
Neutral	-0.44	(2.12)	[18]	0.22	(1.80)	[18]	-0.11	(1.97)	[36]
Overall	-0.41	(2.24)	[34]	1.00	(1.86)	[37]	0.32	(2.16)	[71]

Panel B: Two-Way ANOVA Tests of Between-Subjects Effects

Source	Sum of Square	df	Mean Square	F-statistic	p-value
Readability	34.12	1	34.12	8.50	0.00
Language Sentiment	11.09	1	11.09	2.76	0.05*
Readability * Language Sentiment	9.23	1	9.23	2.30	0.07*
Error	268.99	67	4.02		

Panel C: Contrast Testing for H1

H1: $3 \times$ Positive Language/Low Readability + Neutral Language/Low Readability - $2 \times$ Positive Language/High Readability - $2 \times$ Neutral Language/High Readability = 0

Source	df	Mean Square	F-statistic	p-value
Contrast (+3, +1, -2, -2)	1	28.43	7.08	0.01

Panel D: Simple Main Effects Tests

Simple Effects	df	Mean Square	F-statistic	p-value
Language Sentiment at High Readability Level	1	0.04	0.01	0.92
Language Sentiment at Low Readability Level	1	21.21	5.28	0.03
Readability at Positive Language Level	1	38.74	9.65	0.00
Readability at Neutral Language Level	1	4.00	1.00	0.32

TABLE 3

More Sophisticated Participants' Earnings Judgments

DV = Future Earnings Performance Prediction

To what extent do you agree that W&Z's **earnings performance will be strong** in the near future?

-5	-4	-3	-2	-1	0	1	2	3	4	5

Strongly DISAGREE   Neutral   Strongly AGREE

Panel A: Descriptive Statistics – Mean (Standard Deviation) [Sample Size]

Language Sentiment	Readability								
	High			Low			Overall		
Positive	0.32	(1.70)	[19]	-0.88	(2. 15)	[17]	-0.25	(1.99)	[36]
Neutral	0.22	(2. 29)	[18]	1.41	(2.35)	[17]	0.80	(2.36)	[35]
Overall	0.27	(1.98)	[37]	0.26	(2.50)	[34]	0.27	(2.23)	[71]

Panel B: Two-Way ANOVA Tests of Between-Subjects Effects

Source	Sum of Square	df	Mean Square	F-statistic	p-value
Readability	0.00	1	0.00	0.00	0.99
Language Sentiment	21.44	1	21.44	4.74	0.02*
Readability* Language Sentiment	25.24	1	25.24	5.58	0.01*
Error	303.10	67	4.52		

Panel C: Simple Main Effects Tests

| Simple Effects	df	Mean Square	F-statistic	p-value
Language Sentiment at High Readability Level	1	0.08	0.02	0.89
Language Sentiment at Low Readability Level	1	44.74	9.89	0.00
Readability at Positive Language Level	1	12.88	2.85	0.10
Readability at Neutral Language Level	1	12.37	2.74	0.10

Note: One-tailed p-values are noted with *.

Figure 1: Future Earnings Performance Evaluation - Less Sophisticated Investors

Figure 2: Future Earnings Performance Evaluation - More Sophisticated Investors

- high readability
- low readability

1.41

0.22

0.32

-0.88

Neutral Language   Positive Language

Earnings Performance in the future
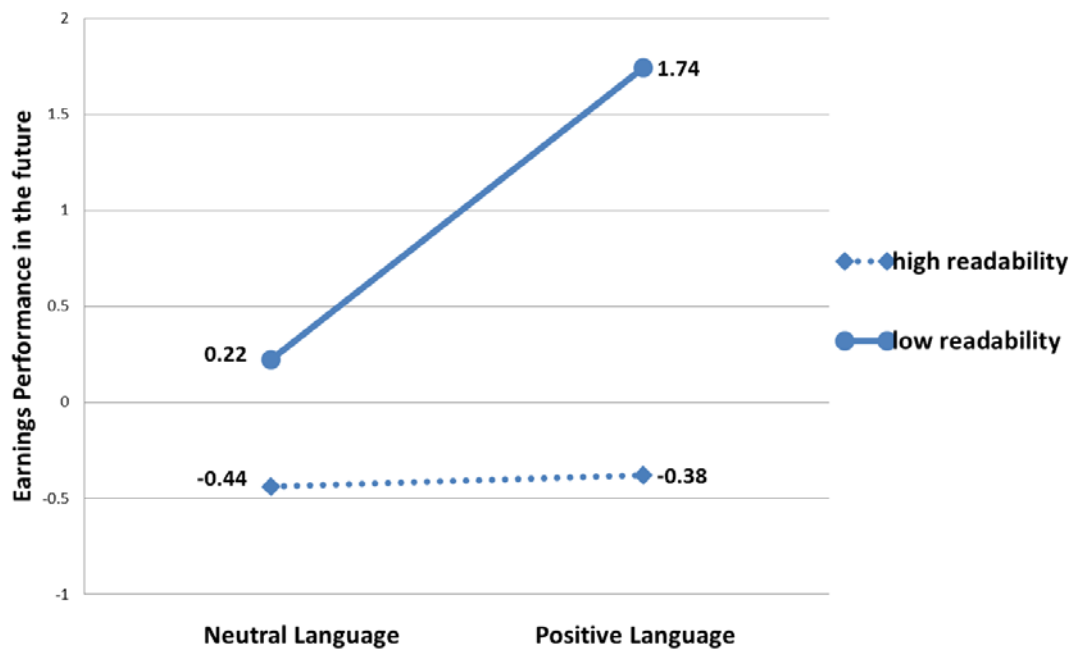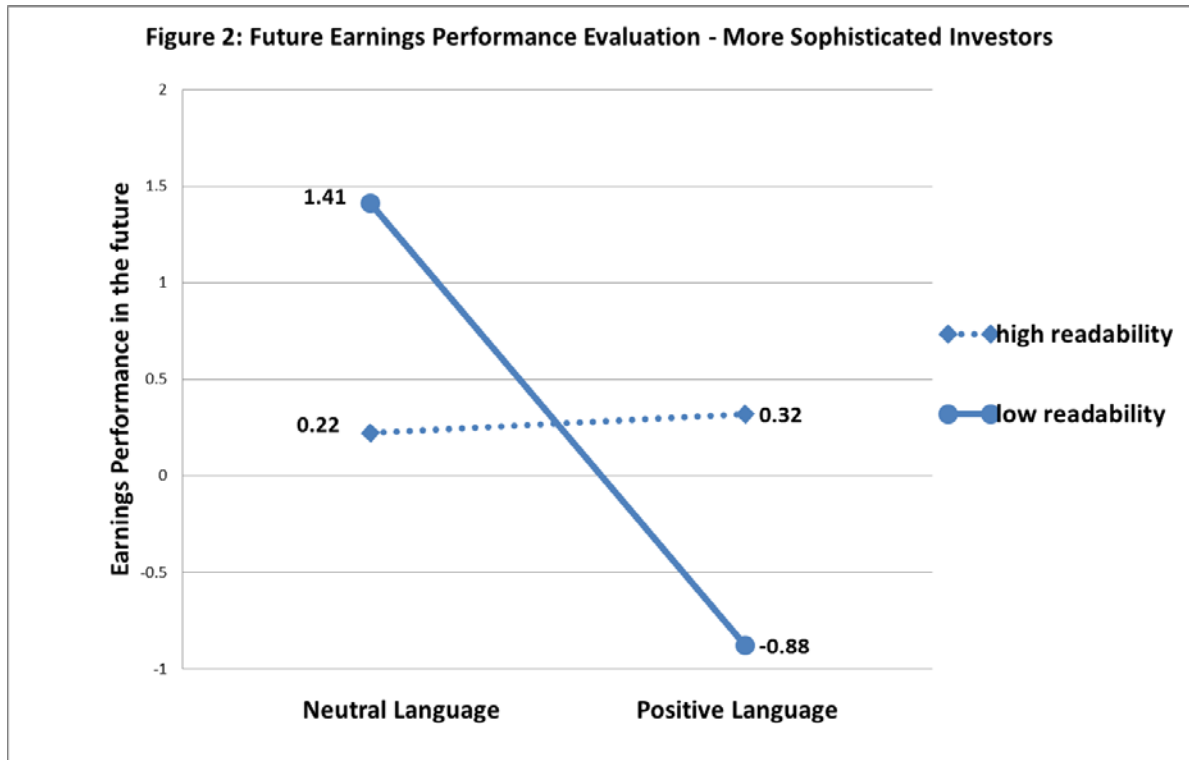
Figure 3: The Mediating Role of Understanding for Less Sophisticated Investors
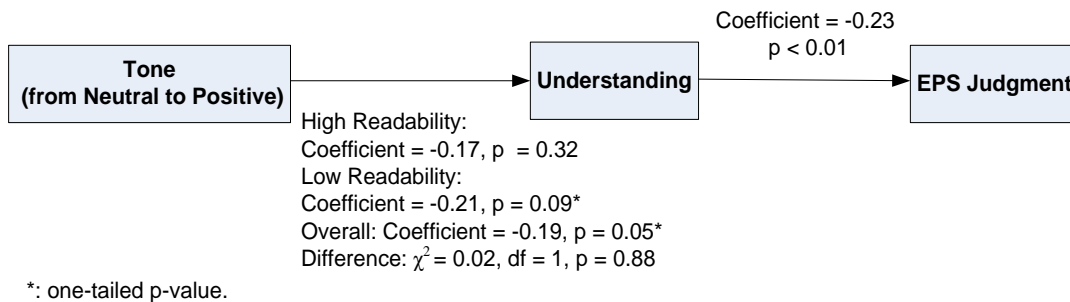
| Tone (from Neutral to Positive) | → | Understanding | → | EPS Judgment |

Coefficient = -0.23
p < 0.01

High Readability:
Coefficient = -0.17, p = 0.32
Low Readability:
Coefficient = -0.21, p = 0.09*
Overall: Coefficient = -0.19, p = 0.05*
Difference: $\chi^2 = 0.02$, df = 1, p = 0.88

*: one-tailed p-value.

This figure shows structural-equation modeling results for the mediating role of understanding for less sophisticated investors. We present the standardized coefficients and corresponding p-values next to each link. Overall model fit indices are as follows: (a) $\chi^2 = 8.18$, df = 6, and p = 0.32, (b) Incremental Fit Index (IFI) = 0.92, which is above the cutoff of 0.90 for model acceptance (Byrne [2001]), and (c) Root Mean Square Error of Approximation (RMSEA) value = 0.04, below the cut-off point 0.08 for good fit (MacCallum, Browne, and Sugawara [1996]).

Figure 4: The Mediating Role of Credibility for More Sophisticated Investors
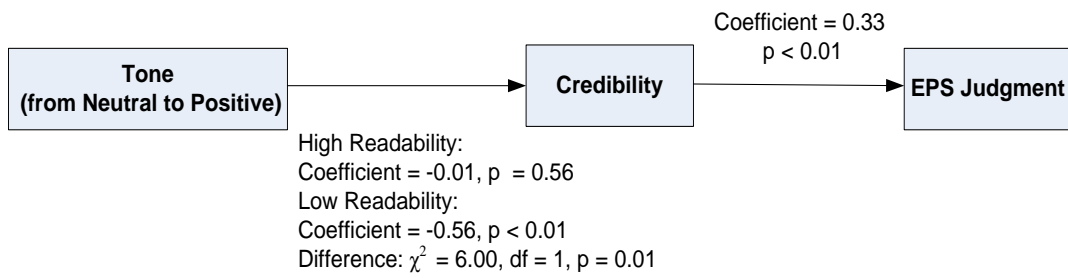
Coefficient = 0.33
p < 0.01

Tone
(from Neutral to Positive) → Credibility → EPS Judgment

High Readability:
Coefficient = -0.01, p = 0.56
Low Readability:
Coefficient = -0.56, p < 0.01
Difference: χ^2 = 6.00, df = 1, p = 0.01

This figure shows structural-equation modeling results for the mediating role of understanding for less sophisticated investors. We present the standardized coefficients and corresponding p-values next to each link. Overall model fit indices are as follows: (a) χ^2 = 5.79, df = 5, and p = 0.33, (b) Incremental Fit Index (IFI) = 0.98, which is above the cutoff of 0.95 for a well-fit model (Byrne [2001]), and (c) Root Mean Square Error of Approximation (RMSEA) value = 0.03, below the cut-off point 0.08 for good fit (MacCallum, Browne, and Sugawara [1996]).