Analyzing Speech to Detect Financial Misreporting

Jessen L. Hobson*
William J. Mayew†
Mohan Venkatachalam†

*Department of Accountancy, University of Illinois at Urbana-Champaign
† Duke University – Fuqua School of Business

Forthcoming, *Journal of Accounting Research*

Abstract

We examine whether vocal markers of cognitive dissonance are useful for detecting financial misreporting. We use speech samples of CEOs during earnings conference calls and generate vocal dissonance markers using automated vocal emotion analysis software. We begin by assessing construct validity for the software-generated dissonance markers by correlating them with four dissonance-from-misreporting proxies obtained in a laboratory setting. We find a positive association between these proxies and vocal dissonance markers generated by the software, suggesting the software's dissonance markers have construct validity. Applying the software to CEO speech, we find that vocal dissonance markers are positively associated with the likelihood of irregularity restatements. The diagnostic accuracy levels are 11% better than chance and of similar magnitude to models based solely on financial accounting information. Moreover, the association between vocal dissonance markers and irregularity restatements holds even after controlling for financial accounting and linguistic based predictors. Our results provide new evidence on the role of vocal cues in detecting financial misreporting.

We thank Dan Ariely, Bob Ashton, Bill Baber, Rob Bloomfield, Larry Brown, Brooke Elliott, Merle Erickson (Editor), Kevin Jackson, Rick Larrick, Feng Li (Discussant), Bob Libby, Charles Lee, Russ Lundholm, Laureen Maines, Maureen McNichols, Mark Nelson, Chris Parsons, Mark Peecher, Madhav Rajan, Doug Stevens, Greg Waymire, Sarah Zechman, Mark Zimbelman, and an anonymous referee for helpful comments and discussions. We also appreciate suggestions from workshop participants at the 2009 BYU Accounting Research Symposium, University of British Columbia, University of Chicago, Cornell University, Florida State University, George Washington University, Georgetown University, Georgia State University, Harvard University, Indiana University, 2011 Journal of Accounting Research Conference, 2010 AAA Annual Meeting, Lancaster University, Oklahoma State University, Queen's University and Stanford University. Ozlem Arikan, Zhenhua Chen, Virginia Chung, Katie French, Mickey Hartz, Chris Kalogeropoulos, Woo Chang Kim, Hui Liu, Hongling Ma, John Montgomery, Kristina Moultrie, Ryan Sturkoff, Simeon Tzolov, Belinda Wen, Christina Woytalewicz, and Yuepin Zhou provided excellent research assistance. The authors declare no financial interests in the manufacturer of the vocal emotion analysis software used in the study.

Analyzing Speech to Detect Financial Misreporting

1. Introduction

In this study, we examine whether nonverbal vocal cues elicited from speech are useful in detecting intentional deception in financial reporting. Detecting deceptive financial reporting is an increasingly important concern for auditors, regulators, investors, and the various constituents that interact with corporations. High profile accounting scandals such as Enron, WorldCom, Tyco, and Satyam have cost market participants several billions of dollars and eroded confidence in published financial statements. These events call into question the ability to uncover financial misstatements by auditors who review and provide an opinion on the financial statements (PCAOB [2007], [2010]). Even sophisticated market participants such as institutional investors and analysts have been remarkably unsuccessful at detecting financial fraud (Dyck et al. [2010]). Therefore, developing a framework to predict deceptive financial reporting in a timely manner can be particularly useful to investors, auditors, analysts and regulators.

Research in social psychology (e.g., Zuckerman and Driver [1985]; Vrij et al. [2000]; DePaulo et al. [2003]) suggests that deceivers' emotions and cognitive processes likely result in many different markers that can help identify deception, such as verbal linguistic cues (e.g., speech content), nonverbal cues (e.g., tone of voice, facial expressions or gestures), and physiological changes (e.g., heart rate). While prior research finds some support that each of these classes of markers help detect deception, Bond and DePaulo [2006] show that individuals (even experts) display only modest accuracy in correctly identifying deceit (see also Vrij [2008]). Part of the challenge that individuals face in detecting deception is the identification of behavioral cues (markers) associated with deception.

The existence and quality of automated tools for identifying deception markers from word usage (verbal cues) has advanced substantially in recent years. Using Linguistic Inquiry and Word Count (LIWC) software, Newman et al. [2003] document a set of "lying words" that identify deceptive language in a variety of experimental settings. Numerous studies in finance and accounting have now applied similar linguistic analysis to various corporate texts in order to detect financial misreporting and to assess fraud risk (see for example Burns et al. [2010], Loughran and McDonald [2011a], Loughran and McDonald [2011b], Larcker and Zakolyukina [2011], Purda and Skillicorn [2010] and Humpherys et al. [2011]).[1] The explosive growth in this area of research stems from both the increased availability of software programs to systematically code language and easy accessibility of large volumes of corporate text files. The general finding from this body of work is that across different software programs (e.g. LIWC, custom dictionaries, Naïve Bayes learning models) and different corporate texts (e.g. 10-K MD&A, earnings conference calls), linguistic deception markers have some predictive ability.

We extend this line of research by examining the predictive ability of nonverbal deception cues, in particular vocal cues, for financial misreporting. Vrij et al. [2000] suggest that verbal and nonverbal measures are complementary mechanisms in detecting deception, but little is known about the information contained in nonverbal cues in the capital market setting. Experimental work by Elliott et al. [2011] finds that communication of restatements via online video, a venue containing both vocal and visual nonverbal cues, impacts investors' perceptions of managerial trustworthiness and investment decisions. Archival work by Mayew and Venkatachalam [2012] suggests that vocal emotion cues exhibited by managers during earnings conference calls have information content. In particular, they document that positive (negative) affective states elicited from voice are positively (negatively) associated with contemporaneous stock returns and future profitability. However, they do not establish why negative affective states arise in the first place.[2] In this paper, we identify one particular reason for CEOs to exhibit negative affective states during conference calls – financial misreporting.

We focus on one vocal marker of deception, cognitive dissonance. Cognitive dissonance is a state of psychological arousal and discomfort occurring when an individual takes actions that contrast

---

[1] A much broader set of earlier research more generally establishes the informational role of linguistic cues. For research that uses newspaper articles see Kothari et al. [2009], Tetlock [2007] and Tetlock et al. [2008]. For company press releases, see Demers and Vega [2008], Engelberg [2008], Henry [2008], Davis et al. [2008]. For earnings conference calls see Matsumoto et al. [2011]. For SEC filings and IPO prospectus, see Li [2008], [2010], and Feldman et al. [2010].

[2] Mayew and Venkatachalam [2011] document a negative association between negative affective states and future stock returns, which is potentially consistent with investors learning about negative news events in the future. However, their study is silent on the specific negative event (such as an adverse financial restatement) that evokes the CEOs' negative affective state.

with a belief, such as cheating while believing oneself to be honest (Mazar et al. [2008]; Graham [2007]). Consistent with Mazar et al. [2008], we posit that, on average, individuals who engage in deceptive financial reporting will experience negative emotions due to cognitive dissonance (Festinger [1957]; Festinger and Carlsmith [1959]). To provide evidence on the predictive ability of vocal cognitive dissonance markers for financial misreporting, we obtain speech samples of CEOs from their interactions with analysts and investors during earnings conference calls. We then measure cognitive dissonance in CEO speech using an automated vocal emotion analysis software based on Layered Voice Analysis (LVA) technology.

While there has been explosive growth in the number of studies using automated vocal emotion recognition in the last decade, no tool or set of acoustic features has emerged as a standard, particularly for analyzing real world speech samples (Schuller et al. [2011]; Patel et al. [2011]; Schuller [2010]). The LVA technology is an emerging and unproven commercial tool that has been sparingly used in the financial markets domain. As such, before examining the predictive ability of vocal dissonance markers in CEO speech, we first assess the construct validity of the cognitive dissonance metric produced by the LVA software.

In a laboratory setting, we adapt the research design in Mazar et al. [2008] to i) generate a sample of truth-telling and misreporting subjects and ii) invoke cognitive dissonance in the misreporting subjects. We then assess whether the LVA dissonance markers extracted from subject interviews are associated with four dissonance-from-misreporting proxies: belief revisions, confessions of misreporting, unexpected reported scores, and a factor score that combines these three individual proxies. We find that the LVA based vocal dissonance markers extracted from the early portion of the interview – which is precisely when cognitive dissonance should be most salient – are increasing in each of these four proxies, both in a univariate setting and after controlling for subject demographics. This collective evidence from laboratory-generated data is consistent with the LVA software capturing dissonance from misreporting.

We then proceed with our archival analysis and investigate whether CEOs that exhibit dissonance during conference calls are more likely to be deceptive with respect to the financial statements. We identify deceptive financial reporting *ex post* as those financial statements adversely restated in the future due to an irregularity. Adverse irregularity restatements represent settings in which managers are most likely to have intentionally misreported (overstated) the firm's financial position. Logistic regressions reveal a positive association between vocal dissonance markers and adverse irregularity restatements. Diagnostic accuracy levels are 11% better than chance and of similar magnitude to models based on financial statement predictors of adverse irregularity restatements. We also find vocal dissonance markers to be associated with irregularity restatements even after controlling for financial statement based predictors and other factors including linguistic deception markers. Overall, we conclude that analyzing speech is useful for assessing the likelihood of financial misreporting.

This study contributes to accounting literature and practice in several ways. First, it provides evidence that elements of voiced speech can help detect financial misreporting. In our use of vocal emotion analysis software, we add a new tool to the current approach of predicting financial misstatements that is almost exclusively based on quantitative financial measures or linguistic features. We acknowledge that identifying deception using speech is part of a large body of research on detecting deception (see DePaulo et al. [2003] for a summary), and we are not the first to undertake such an endeavor. However, to our knowledge, this paper is the first to investigate the predictive ability of vocal cues for actual misreporting in a capital market setting using ecologically valid speech samples from a large cross section of corporate executives. In this way, we begin to fill the void noted by Hirschberg [2010] regarding a relative lack of evidence on vocal cues for deception detection.

Second, our evidence suggests that investors, analysts, auditors and other parties that rely on communications with management pay particular attention to both the questions they ask of the management and the answers that they receive. In other words, our evidence highlights the importance of interactions between executives and investors during earnings conference calls, road shows, financial press appearances, shareholder presentations, etc. In addition, auditors responsible for attesting to a firm's financial statements could potentially consider speech analysis as an additional input into their assessment of fraud risk. The Public Company Accounting Oversight Board (PCAOB) has recently emphasized the importance of examining quarterly earnings calls as a means of detecting fraud (PCAOB [2010]). Our research provides evidence in support of this emphasis.

Third, this paper adds to prior work (Mayew and Venkatachalam [2012]; Han and Nunes [2010]) that uses LVA based software as a tool to measure vocal emotions by providing construct validity for one of the measures produced by the software. However, a limitation of this software is that the precise inner workings are proprietary. As such, to begin to understand the precise vocal features that mark dissonance or indicate misreporting in financial markets, future research should also consider the use of alternative software tools such as Praat (Boersma and Weenink [2010]; Owren [2008]). Moreover, vocal features represent only one class of deception markers (DePaulo et al. [2003]; Stice [1992]) and an evaluation of alternative nonverbal markers of deception is a worthwhile endeavor. We view our results as a starting point for future research that considers properties of managerial communication beyond financial accounting numbers for the detection of misreporting in capital markets.

2. Prior Research and Hypothesis Development

Deception is a deliberate attempt to mislead (DePaulo et al. [2003]). Financial misreporting is a particular type of deception intended to deceive a company's stakeholders. Prior archival work in financial accounting has primarily explored the predictive ability of financial measures (Beneish [1997]; Dechow et al. [2011]) and nonfinancial performance measures (Brazel et al. [2009]) in detecting financial misreporting and assessing fraud risk.[3] While quantitative information contained in the financial statements represents a significant component of the overall communication of a firm's strategic decisions and outcomes, managers supplement mandated disclosures with press releases, conference presentations and earnings conference calls to elaborate on previously disclosed information or provide timely new information to market participants and contracting parties. With the passing of Regulation FD in 2000 and subsequent innovations in technology over the last decade, vast audiences can access corporate

[3] Experimental work in managerial accounting has also investigated misreporting, primarily with a focus on how the extent to which management control systems reduce lying and/or promote honesty. Salterio and Webb [2006] review this literature. Our investigation differs from this stream of literature in that we focus on detecting misreporting as opposed to designing a control system that prevents misreporting.

communications in the form of audio, and more recently, video broadcasts. Such communications, particularly spontaneous and interactive communications of firms with analysts during earnings conference calls, provide market participants with a richer information set that includes both verbal and nonverbal cues.

Market participants and regulators seem to believe that verbal and nonverbal cues can help identify misreporting and assess fraud risk. Equity research firms employ former CIA (Central Intelligence Agency) agents to search for verbal and nonverbal cues of deception during corporate earnings conference calls (Javers, [2010]). The Public Company Accounting Oversight Board (PCAOB) recently issued Auditing Standard No. 12, which explicitly mandates that auditors consider "observing or reading transcripts of earnings calls" as part of the process for identifying and assessing risks of material misstatement (PCAOB [2010]). Unfortunately, the precise verbal and nonverbal cues identified and used by equity research firms are proprietary, and the authoritative auditing standards are silent on what specifically an auditor should "observe" during an earnings call. To begin to fill this void, we examine whether vocal markers of cognitive dissonance can assist in predicting financial misreporting.

For both practical and theoretical reasons we focus on one particular vocal marker, cognitive dissonance. Research in psychology (e.g., Zuckerman, DePaulo and Rosenthal [1981], Horvath [1979]) suggests several potential ways to discriminate deceivers from truth tellers, including physiological traits (e.g., blood pressure, heart rate, brain activity), kinesics (e.g., facial expressions, body movements), word usage (lexical features), and vocal profiles. We focus on nonverbal vocal deception markers, as opposed to other deception markers, for several reasons. First, capturing physiological changes of corporate executives during earnings conference calls via methods such as brain fMRI (functional magnetic resonance imaging) or skin conductance tests is impractical. Second, verbal linguistic markers of deception in the financial reporting context is the subject of numerous recent studies (Burns et al. [2010]; Loughran and McDonald [2011a], [2011b]; Larcker and Zakolyukina [2011]; Purda and Skillicorn [2011]; Humpherys et al. [2011]). Third, while software programs can measure changes in facial expression (Meservy et al. [2005]; Jensen et al. [2008]), automated kinesics is still in a state of infancy, and video broadcasts of corporate earnings are still quite rare. Finally, audio broadcasts of earnings conference calls are now common and commercial software products to process audio are emerging.

Conditional on our interest in vocal markers, we then focus on cognitive dissonance for the following reasons. First, research in psychology suggests that emotions are conveyed through an individual's voice (Juslin and Laukka [2003]; Juslin and Scherer [2005]; Scherer [1986]), and deceivers commonly experience fear, anxiety, guilt, and shame. Some of these emotions stem from the prospect of being caught, while other emotions stem from cognitive dissonance. Cognitive dissonance is a feeling of psychological discomfort felt when one's actions and beliefs are discrepant (Festinger [1957]). Therefore, deceivers are expected to experience more cognitively dissonant feelings. Second, Javers [2010] notes that former CIA agents hired by equity research firms search earnings conference calls specifically for markers of cognitive dissonance. Third, recent experimental research by Mazar et al. [2008] directly links misreporting and cognitive dissonance. Mazar et al. [2008] discusses the aversive feeling individuals experience during or after cheating and argue that this aversion results because individuals generally want to view themselves as being honest. They find that, in a setting where subjects are given incentives to cheat by misreporting performance for personal gain, simple reminders of the emotional costs of deviating from the self-concept of honesty (i.e. cognitive dissonance costs) substantially dampen individuals' propensities to misreport. Finally, from a practical standpoint, we have access to a commercial software product that purports to capture emotions related to cognitive dissonance.

Based on the discussion above, we hypothesize:

H1: The probability of financial misreporting is positively associated with the extent of cognitive dissonance markers contained in the CEO's voice.

Our generic empirical design to test for the association predicted by H1 entails i) obtaining speech samples from a distribution of misreporters and truth-tellers, ii) measuring the level of cognitive dissonance contained in the vocal wave for each observation, and iii) assessing the predictive ability of dissonance markers for misreporting by estimating a logistic regression of the following form:

$$\Pr(\textit{Misreporting}) = f(\textit{Vocal Dissonance Markers}) \qquad (1)$$

Two conditions must hold for us to observe valid evidence consistent with H1: (1) misreporters must feel cognitive dissonance when misreporting, and (2) our vocal dissonance markers must capture cognitive dissonance stemming from misreporting without significant measurement error. Regarding the first condition, if corporate executives are inherently overconfident, they may never believe they are misreporting and in turn may not experience dissonant feelings. On the other hand, former Satyam Chairman B. Ramalinga Raju, in his letter admitting fraud, stated that he was carrying a "tremendous burden on his conscience." That a CEO would reference a burden on his conscience suggests that dissonance may occur even at the highest levels of corporate management. However, *ex ante* one might challenge that our predictions under H1 might not hold in an empirical archival setting.

Regarding the second condition, we use dissonance markers produced by the LVA software (described in more detail in section 3.2). We are unaware of any systematic archival or experimental evidence that directly validates LVA's ability to capture cognitive dissonance from free flowing speech.[4] Mayew and Venkatachalam [2012] present indirect archival evidence by presenting a negative relation between stock returns and LVA based cognitive dissonance measures extracted from CEO dialogs with financial analysts. While this evidence is potentially consistent with the software capturing cognitive dissonance, it does not speak to cognitive dissonance arising from misreporting. Moreover, although corporate executives may feel dissonant when misreporting, through speech training they may be able to mask their emotions during conference calls in a manner undetectable by the software.

3. Construct Validity of Vocal Dissonance Markers

Before we proceed to test H1 in the archival setting, it is important to assess whether the LVA dissonance measure captures the construct of interest, which is cognitive dissonance from misreporting.

[4] Gamer et al. [2006] experimentally investigate LVA based cognitive dissonance levels as part of an overall assessment of all LVA metrics provided in an early version of the LVA software and find that LVA outputs are statistically indistinguishable across liars and truth tellers. Audio files constructed in Gamer et al. [2006] restricts experimental subjects to monosyllable verbal responses of "yes" and "no" to questions rather than free flowing responses, which may have significantly reduced the predictive power of the metrics generated by the LVA software (Palmatier [2005]). See Mayew and Venkatachalam [2011] for a literature review of the studies investigating LVA based metrics.

The ideal research design for establishing construct validity would be to observe and measure variation in the extent of dissonance felt from misreporting among subjects, extract speech samples from each individual, and correlate the LVA dissonance markers measured from the speech samples with a measure of the extent to which the speaker felt dissonance from misreporting. Directly measuring the extent of dissonance from misreporting in a speaker is theoretically possible with brain imaging technology (Harmon-Jones [2004]). However, for practical considerations we must rely on indirect measures. A laboratory setting is ideal for assessing construct validity because we can (1) construct multiple indirect measures of the construct of interest, (2) generate speech samples in an environment that is less susceptible to extraneous background noise, and (3) follow protocols from existing studies to generate variation in misreporting and cognitive dissonance in subjects.

3.1 Generating variation in dissonance from misreporting

Our first research design choice pertains to generating a distribution of truth tellers and misreporters. We provide monetary incentives in a manner that allows subjects to choose between misreporting and truth telling. That is, we allow subjects to endogenously choose to misreport (or not) for personal gain (e.g., Evans et al. [2001]; Hales et al. [2011]; Mazar et al. [2008]). To generate dissonance in subjects choosing to misreport, we use a design feature of Mazar et al. [2008], who show that binding cognitive dissonance costs can be generated in a laboratory setting. In particular, they show that it is possible to increase the emotional costs of deviating from personal honesty norms (i.e., costs from cognitive dissonance) by reminding subjects of their personal moral codes via subject recitation of the Ten Commandments. We capitalize on this feature in our design by first incentivizing misreporting for personal gain to generate an endogenous sample of truth tellers and misreporters. Then, we infuse reminders of moral codes *after* participants have reported their score on a private task. The purpose of this timing of the moral code reminder is to stimulate the emotional burden of cognitive dissonance in misreporters, the emotional markers of which the LVA software claims to capture from voice. We describe the research design in more detail below.

3.1.1 Design Timeline and Procedure

Fifty-nine undergraduate volunteers from two large public U.S. universities participated in a two-part laboratory session (see timeline of events in Figure 1). Participants were 37% female, with a median age of 20 years, in their sophomore year, and had completed three (one) math (English) college courses (see Panel A of Table 1). The first part of the study was an online portion containing initial, general instructions, Scholastic Aptitude Test (SAT) background instructions and examples, and a self-timed, five-minute SAT test.[5] Participants were given 4 points for each correct answer, -1 point for each incorrect answer, and 0 points for each skipped question. Responses were graded automatically through the online interface, which revealed to the participant how many questions had been answered correctly (labeled *SURVEY*). After receiving this feedback, subjects were asked to predict how many SAT questions they could answer correctly if they were to take the SAT test again using similar questions. The answer to this question captures the subjects' beliefs about their ability to answer SAT questions before entering the laboratory and we label this *BELPRE*. As a whole, the purpose of this online portion was to i) re-acquaint the student with SAT questions, and ii) initiate a prediction of self-assessed ability. Panel A of Table 1 reveals that average subjects scored 11.63 points (*SURVEY* = 11.63) and believed they could answer 6.00 SAT questions correctly if given an additional 5 minutes.

After completing the online portion, participants completed the laboratory portion.[6] The laboratory portion consisted of four activities: (1) taking, scoring and reporting the results from a timed, five-minute SAT test, (2) filling in answers to a set of questions on a midpoint questionnaire, (3) answering a set of interview questions while being video recorded, filling in answers to an exit questionnaire, predicting performance on a future SAT test, and finally, (4) being paid and debriefed.

[5] At one of the universities, students traditionally take the ACT exam to qualify for admission. As such, at that university, we labeled all materials ACT instead of SAT.

[6] Aside from answering the timed SAT questions, participants were able to complete the online portion of the experiment at their own time and pace. We made sure that participants had finished the online portion before starting the laboratory portion, and that they completed the online portion only once. The duration between online completion and the laboratory portion ranged from 14 days to one hour (median number of days equals 1). The number of days is not significantly correlated with participants' scores on either the laboratory SAT questions or the online SAT questions.

The laboratory portion proceeds as follows. First, after some brief initial instructions by a student administrator, participants took a timed, five-minute SAT test. This test had questions that were similar to those in the online portion of the study. Next, the participants self-graded their answers and reported an overall score on a separate score sheet, which determined their payoff. Participants were informed they would be permitted to retain their test and answer sheets and only needed to hand in a sheet containing their reported score for determining payoffs. The student administrator left the room both when the SAT test was taken and when the SAT test was self-scored. During this time, the participants had both the test form and the answer sheet. The purpose of informing participants that they would not be turning in their original testing sheets, of having the student administrator leave the room, and for using a student for administration instead of one of the paper's authors was to lower perceptions of monitoring, and in turn invoke misreporting in subjects. Additionally, allowing subjects to retain their test sheets was meant to prevent subjects from feeling fearful of being covertly detected. Fear of detection is a reasonable emotion for a misreporter facing a non-zero probability of detection, but fear is not the emotion of interest in this study.

Second, participants answered a midpoint questionnaire. The purpose of this questionnaire was to obtain demographic information and make the participants cognizant of their own personal moral code. To this end, we asked all participants to write down as many of the Ten Commandments as they could remember. Participants are likely to be aware that the Ten Commandments represent a moral code regardless of their personal religious beliefs (Mazar et al. [2008]).[7] This moral code reminder was intended to invoke emotions associated with cognitive dissonance in misreporters.

Next, a laboratory administrator separately videotaped each participant's answers to interview questions in a separate interview room. All instructions and interview questions were prerecorded and sequenced with a PowerPoint presentation. The administrator operated the video equipment, played and advanced the prerecorded audio in the PowerPoint, and prompted the interviewee to expand their answers

[7] We find that the Ten Commandments were widely known to our subject pool. On average participants correctly recalled five of the Ten Commandments. More than 90% of participants correctly recalled at least two of the Ten Commandments.

if the answers were overly short. Thus, the administrator's interaction with each participant was minimal and the interviewer did not alter interview questions in a strategic manner. Naturally, in real-world settings, question and answer dialogs between corporate executives and analysts are likely strategic and dynamic, with subsequent analyst questions being conditioned on answers to preceding questions. While our design lacks this realism, we crafted interview questions to parallel the type of questions commonly asked in an earnings conference call. The prerecorded interview had seven questions. The first question was innocuous and calibrated the participant's voice for the vocal emotion analysis software. The remaining six questions pertained to reported performance of the participant, with questions ranging from general to specific, similar to the progression of questions during earnings conference calls.[8] For example, the first question asked the subject to verbally repeat the score they reported on the score sheet (much as managers repeat reported earnings from the press release when beginning discussions on a quarterly earnings conference call). The second question asked whether the reported score was better or worse than expected, and the third question asked about what the most difficult portions of the test were and why. Importantly, no question directly asked whether the subject was able to achieve their reported score via misreporting, and as such, we never explicitly solicited for participants to make an untrue statement. This is an externally valid feature of our design, in the sense that in public capital market settings, executives are rarely, if at all, asked questions about misreporting *ex ante* by financial analysts. However, it is likely that questions center around issues where analysts have some suspicion of misreporting.

[8] The calibration question is as follows: "To help us calibrate our equipment and make sure we are ready, please do the following three things: Describe the room you are sitting in. Spell the following words letter for letter: Dictionary and Abbreviation. Read the following numbers aloud: 1,965; 818; 11,757;" The six interview questions were as follows. (1) "First, please restate the score that you wrote down on the answer sheet." (2) "Was your performance on the SAT questions you just answered better or worse than how you have done in the past? Please explain your answer in detail." (3) "Which types of SAT questions did you find most difficult to answer? Why were these questions so difficult? What strategies did you use to answer these difficult questions? Please explain in detail." (4) "Many of our participants score below 10 points on these SAT questions. Describe as completely as possible how you were able to achieve the score you reported." (5) "Overall, how do you feel about the SAT score you just reported? Please explain in detail." (6) "How would you respond to someone that told you they thought the SAT score you just reported was too high? Please explain your answer in detail."

Subsequent to the interview, participants were taken back to the original room to answer a final questionnaire containing additional demographic and manipulation check questions. Also, participants predicted the number of SAT questions they could answer correctly in a hypothetical future session, the answer to which we label *BELPOST*. This prediction question had the exact same wording as the prediction question at the end of the online survey (*BELPRE*). The difference between *BELPOST* and *BELPRE* captures the changes in a participant's belief about his/her ability to successfully answer SAT questions. We label this difference *BELREV*. This variable is similar to the attitude-change measure commonly derived in the induced compliance cognitive dissonance paradigm (Cooper [2007]; Harmon-Jones and Mills [1999]; Elliott and Devine [1994]). One way participants can resolve the cognitive dissonance resulting from overstating their SAT score is to modify their belief about how competent they actually are at answering SAT questions. Being unable to change their misreporting behavior after reporting, participants may instead modify their beliefs in order to help resolve their cognitive dissonance. Thus, we would expect that participants who have overstated their SAT score would have a relatively higher second prediction score, and hence higher values of *BELREV*. *BELREV* serves as our first proxy for the extent of dissonance from misreporting.

Finally, participants were individually taken to a separate room to be paid and debriefed. Participants were paid \$5 for completing the online survey, and \$10 for coming to the laboratory session. As mentioned before, each participant was compensated based upon the *self-reported* number of points scored on the SAT test. The average participant reported scoring 24.69 points and each participant received \$0.50 for each point, yielding an average payout of \$12.35. Thus, participants earned \$27.35 on average. In addition, all participants were entered into a random drawing for a \$500 prize.

The laboratory portion took an average of seventy-five minutes, with the video recording taking an average of five minutes. After receiving their payment, participants were informed of the research purpose and that the researchers had intended that some participants would overstate their true SAT score. Once assured that overstatement was acceptable and that no punishments for overreporting would accrue, participants were asked whether they had overstated their score. Response to this question formed our second proxy for dissonance from misreporting, *CONFESS*, which is coded as one if the participant admitted to overstating the reported SAT score and zero otherwise. This opportunity to confess serves as another opportunity for the participant to resolve dissonance (Stice [1992]) and provides an alternative way to identify participants experiencing cognitively dissonant feelings from misreporting. Providing this debriefing information to the subjects after payment was a mandatory condition in our design to ensure that the subjects were not harmed from our infliction of cognitive dissonance.

Our third proxy for the extent of dissonance from misreporting, *USCORE*, is the difference between the self-reported score (*SCORE*) and score obtained in the online session (*SURVEY*). This measure is motivated based on the finding in Murphy [2010] that self-reported dissonance scores are increasing in the magnitude of misreporting. In our setting, higher unexpected scores (*USCORE*) proxy for instances in which subjects have misreported more, and should in turn experience more dissonance.

Our experimental design choices preclude us from unambiguously identifying misreporters. Hence, an important caveat with respect to the three proxies of dissonance from misreporting, *BELREV*, *CONFESS* and *USCORE* is that they are imperfect and contain measurement error. However, to the extent we obtain consistent evidence using each of these measures, it is unlikely that any measurement error inherent in these measures is driving our findings.

3.2 Vocal Measurement of Emotions Stemming from Cognitive Dissonance

To generate speech samples for analysis, we replay each video and manually isolate only the audio of each participant's answers to the interview questions.[9] The free flowing speech extracted from each video was encoded in mono directly onto computer hard disk, using Total Recorder 7.1 software, at 11.025 kHz sampling rate and 16-bit quantization, and saved as uncompressed .wav files. The .wav audio files were then analyzed using a commercial version of the LVA software that uses Layered Voice Analysis (LVA) technology developed by Nemesysco Ltd..

LVA is based on a set of proprietary signal processing algorithms purported to identify different types of stress, cognitive processes, and emotional reactions. The algorithms measure features of the

[9] The videos were created with JVC Everio G camcorders.

speech waveform to create a foundation for identifying the speakers' emotional profile.[10] Because waveforms are person specific, the software measures deviations from a calibrated baseline for each speaking subject. Although there are several versions of the LVA software, we use the commercial version developed for business applications called the Ex-Sense Pro R (version 4.3.9). This software has been used in archival work to measure emotion profiles of corporate executives in the capital markets (Mayew and Venkatachalam [2012]).[11]

The Ex-Sense Pro R software produces four "fundamental" voice based measures, labeled Emotional Stress Level, Cognition Level, General Stress Level and Thinking Level.[12] Pertinent to our study is Cognition Level that is purported to measure cognitive dissonance.[13] The software also produces other measures deemed "conclusion" variables (e.g., Lie Stress), which are proprietary combinations of the fundamental LVA measures and are meant to indicate when a speech segment may represent untruthful statements. Because our laboratory setting is specifically designed to evoke cognitive dissonance and in the interview we purposefully do not ask direct questions to which the answers will necessarily be untruthful, we do not consider the other fundamental or conclusion variables produced by the software. Moreover, earlier literature has found little evidence that the built in LVA conclusion variables perform better than chance levels (see Mayew and Venkatachalam [2012] for a summary), while more recent work suggests the more primitive fundamental level LVA variables offer better predictive ability (Elkins [2010]; Elkins and Burgoon [2010]).

[10] As a commercial product, the particular vocal attributes being extracted and combined by LVA are not disclosed. Mayew and Venkatachalam [2011] partially reverse engineer the software and show that five basic acoustic features explain approximately 35% of the variation in the LVA cognitive dissonance metric. However, there is debate in the literature regarding the inner workings of the software (Elkins and Burgoon [2010]; Lacerda [2009]), which in part motivated us to undertake a construct validity exercise.

[11] LVA based software products have been used in other contexts for measuring other emotions such as embarrassment, stress associated with posttraumatic stress disorder, and for detecting deception in experimental and field settings. See Mayew and Venkatachalam [2011] for a review of this literature.

[12] For a detailed description of the various parameters and a more complete discussion of the LVA software, please refer to Mayew and Venkatachalam [2011].

[13] Ex-Sense Pro R user manual states: "Cognition Level reflects a situation when two or more non-complimentary logical processes are "processed" in the brain, for example, a logical conflict between what the mouth is saying and what the brain thinks. This is also referred to as cognitive dissonance (Festinger [1957])."

The software provides output at the voice segment level. A voice segment is a logical portion of continuous voice that may range from one word to a few words (usually one to two seconds). Discussions with the software developer suggest Cognition Level values greater than 120 are indicative of dissonance levels that require attention (see Mayew and Venkatachalam [2012]). Hence, we measure cognitive dissonance, *COGDIS*, as the number of voice segments yielding Cognition Level values greater than 120 divided by the total number of segments. Panel A of Table 1 reveals an average *COGDIS* in our sample of 0.217 and a standard deviation of 0.088, similar to those reported in Mayew and Venkatachalam [2012].

3.3 Construct Validity Tests and Results

To provide construct validity for the *COGDIS* variable, we explore the correlations between *COGDIS* and each of our three proxies for the dissonance construct (i.e. *BELREV, CONFESS,* and *USCORE*). Relying exclusively on any one proxy individually to assess construct validity is problematic because each measures the construct of dissonance from misreporting with error. Therefore, we conduct a factor analysis using principal component estimation to group these three variables. We find one factor with an eigenvalue greater than 1 that we label *DISFACTOR*, which serves as our combined measure of dissonance from misreporting. This factor accounts for 55% of the variance in our sample.

Panel B of Table 1 presents the correlation statistics. We observe a positive association between each of the dissonance proxies, although the positive association between confession (*CONFESS*) and belief revision (*BELREV*) is not statistically significant.[14] Turning to construct validity tests, we find that the LVA based vocal marker of cognitive dissonance, *COGDIS*, is positively associated with the belief revision measure, *BELREV*. Recall that *BELREV* captures the revision in a participant's beliefs in order to resolve dissonant feelings, and is a classic *ex-post* indicator that cognitive dissonance was present (Cooper [2007]). We find that the Spearman rank correlation between *COGDIS* and *BELREV* is positive and statistically significant ($\rho = 0.334$, $p = 0.010$). The Pearson correlation is also positive but marginally

[14] One potential reason for the weak association with confession is that *CONFESS* is a dichotomous variable whereas the other variables are continuous and as such contain more information. Additionally, there may be systematic downward bias in *CONFESS* if individuals who misreported do not confess.

significant in a one tailed test (ρ = 0.192, p = 0.072 one tailed). To probe this relationship further, we note that these two measures differ in one critical aspect. While *COGDIS* is derived from measurements at small intervals throughout the entire speech sample, *BELREV* is measured as the revision in beliefs of participants from the beginning to the end of the laboratory session. The intuition for *BELREV* is that participants attempt to remove the uncomfortable feelings associated with dissonance from misreporting by upgrading their beliefs about their own ability to answer test questions. Prior research does not indicate *when* this belief revision occurs other than finding that the revision has occurred prior to the completion of the laboratory study (e.g., see references in Cooper [2007]). In theory, if the LVA software is capturing emotions associated with cognitive dissonance *as they occur*, voice samples analyzed in between the invocation of dissonance and the point of belief revision should better identify dissonance from misreporting. Empirically, this would imply a stronger association between *COGDIS* and *BELREV* when *COGDIS* is measured between the point of dissonance invocation and belief revision than between belief revision and the end of the interview.

To test this conjecture, we define early (late) vocal based dissonance, *E_COGDIS* (*L_COGDIS*) as *COGDIS* from the first (second) half of the interview. The first half of the interview follows the moral code reminder, which is our invocation of dissonance, and hence is more likely to overlap with the period where cognitive dissonance is most likely to be present. If true, the association between *E_COGDIS* and *BELREV* should be stronger than the association between *L_COGDIS* and *BELREV*. Consistent with this intuition, the Spearman (Pearson) correlation between *E_COGDIS* and *BELREV* is 0.440 (0.320) with a p-value of 0.001 (0.014). In contrast, the respective correlations between *L_COGDIS* and *BELREV* are much lower in magnitude at 0.098 (0.032) and not statistically significant. These results are consistent with the LVA software capturing cognitive dissonance from misreporting as it occurs.

Turning to the remaining dissonance proxies, we find similar evidence as in the case of *BELREV*. Voice based dissonance from the early portion of the speech sample (*E_COGDIS*) is positively and statistically associated with *CONFESS*, *USCORE* and *DISFACTOR*, while voice based dissonance from the latter portion (*L_COGDIS*) exhibits no statistical association with any of these variables. Because these are simply associations resulting from a distribution of subjects who self selected into a misreporting condition or not, it is possible that the reported associations are driven by variables correlated with subject demographics. Therefore, we report in Table 2, the association between *E_COGDIS* and each of the dissonance proxies, after including controls for subject age, year in school, number of math and English classes, ability (captured as the number of points scored during the online portion), and gender.[15] The multiple regression results in Table 2 Columns A-D reveal that dissonance from the early portion of the speech sample is associated with each of the four dissonance from misreporting proxies (*DISFACTOR, BELREV*, *CONFESS* and *USCORE*) after including subject characteristics. Taken together, we view the evidence in Table 1 and Table 2 as suggestive that the voice based dissonance markers generated by LVA have construct validity.

4. Empirical Analysis Using Archival Data

4.1 Design Overview

To test H1, we expand our conceptual prediction model outlined in equation (1) as follows:

Pr(*Misreporting*) = *f*(*Vocal Dissonance Markers, Dissonance Drivers Unrelated to Misreporting, Financial Statement Based Predictors of Misreporting, CEO Characteristics, Monitoring*) (2)

Equation (2) adds four additional conceptual components to the prediction model of misreporting. First, we control for factors other than the act of misreporting that may cause cognitive dissonance in executives. Recall that cognitive dissonance induces negative affect due to a disjoint between beliefs and actions.[16] Suppose CEOs believe they are honest, competent, and in control of their firms. Then, any action by the manager or outcome that is not consonant with these held beliefs would inflict cognitive dissonance. That is, actions that result in poor performance or growth, uncertain firm outcomes and

[15] We exclude voice-based dissonance from the later portion of the speech sample (*L_COGDIS*) because we observe no univariate association with *L_COGDIS* and any of the dependent variables. Including *L_COGDIS* does not impact our inferences with respect to *E_COGDIS* or any other independent variable.

[16] Note that the LVA software measures cognitive dissonance after taking into account the baseline vocal characteristics in the calibration phase of the speech analysis. Therefore, any dissonance that is felt by an executive due to other factors and therefore, inherent in his vocal characteristic, is likely to be differenced away by the LVA software because LVA calibrates each executive's speech for their unique vocal characteristics at the beginning of the speech. Nevertheless, for completeness, we include other drivers of dissonance in the model.

dishonest financial reporting induce dissonance. The LVA software does not however distinguish between sources of cognitive dissonance, but rather measures the extent of dissonance present from whatever source. Mayew and Venkatachalam [2012] provide evidence consistent with higher levels of dissonance in poorly performing firms, and in firms operating in more uncertain environments. Therefore, controlling for these non-misreporting sources of dissonance should yield a more powerful specification. Specifically, we control for performance by including return on assets, unexpected earnings, prior year stock returns and the book to market ratio. We control for uncertain environments with firm size and stock return volatility.

The second addition pertains to existing financial-statement based predictors of misreporting. Since our objective is to assess whether, and to what extent, vocal dissonance markers predict misreporting, adding known predictors of misreporting provides a benchmark against which we can compare the predictive ability of vocal dissonance markers. Moreover, we can assess whether vocal dissonance markers provide incremental predictive ability to financial information. We use two summary metrics from the recent accounting literature to assess the predictive ability of financial statement data in our setting. The first metric is F-Score, developed by Dechow et al. [2011], for the purpose of predicting accounting manipulations disclosed in the SEC Accounting and Auditing Enforcement Releases (AAERs). The second metric is Accounting Risk, a commercially available summary metric, which is purported to capture the likelihood of misrepresentation in corporate financial reports. Recent research shows Accounting Risk is a potent predictor of intentional corporate misstatements (Correia [2010]; Price et al. [2011]).

The third addition pertains to executive characteristics. If older or more seasoned executives are better able to control their emotions and have lower incentives to misreport due to lessened career concerns, we may observe both lower dissonance levels and lower levels of misreporting for experienced executives. That is, the hypothesized positive association between vocal dissonance and misreporting may be driven by executive age and tenure. We therefore control for both of these factors in our empirical specification.

The fourth and final addition considers the role of outside monitors as a potential confounding factor. In the laboratory setting, monitoring was set to a constant low level to help induce misreporting. In archival data, monitoring likely varies in the cross section. Existing research suggests that sell side financial analysts and analysts of higher quality are successful monitors in terms of dampening earnings management (Yu [2008]). We use the incidence of questioning by an all-star analyst during the conference call as our measure of monitoring. If high quality analysts are successful at mitigating earnings management and in turn financial misreporting, firms monitored by all-star analysts should be less likely to misreport. On the other hand, results in Mayew [2008] suggest that managers may endogenously allow or disallow an analyst to participate in an earnings call. In such situations, managers who misreport may have incentives not to allow an all-star analyst to ask a question during the conference call. Thus, our prediction for the relation between the incidence of all-star analyst question and financial restatement is ambiguous.

4.2 Executive voice data

An ideal archival design would compare a sample of executives who are deceptive, and a matched sample of executives in firms with similar economic characteristics but who are not deceptive. Achieving this ideal is difficult for two reasons. First, audio files of executives speaking during earnings conference calls are publicly available for relatively short periods of time, perhaps due to litigation risk concerns (Friedman, [2010]). Although transcripts of the conference calls are available for a large cross section dating back to the passage of regulation FD, the related audio files are re-streamed over the internet for periods typically ranging from one fiscal quarter to one fiscal year from the earnings call date. Data providers such as ThomsonReuters StreetEvents, who provide subscribers with restreaming access for periods specified by the firms, do not allow downloading of the audio files. Researchers, therefore, face enormous data collection costs because the only way to analyze audio files is via re-streaming while publicly available. Adding to the costs, the researcher must then manually isolate and extract the voice of the executive of interest from the conference call dialog in order to conduct a speech analysis for each executive's voice.

Data constraints notwithstanding, a second challenge is that we cannot ensure that managers will discuss a deceptive topic either voluntarily in the presentation or when probed by analysts in the Q&A (Hollander et al. [2009]). We view it as unlikely that managers can avoid speaking about major economic factors affecting their firms altogether, and conjecture that some of the discussion will overlap with topics pertaining to misreporting. To the extent this is not the case, it biases against finding results supporting H1.

Given these data collection challenges, we begin with the sample collected by Mayew and Venkatachalam [2012], where streamed audio from earnings conference calls available on ThomsonReuters StreetEvents was encoded in mono directly onto computer hard disk, using Total Recorder 7.1 software, at 11.025 kHz sampling rate and 16 bit quantization, and saved as uncompressed .wav files. This speech corpus comprises 1,647 quarterly earnings conference calls spanning the period January 1 through December 31, 2007, and represent fiscal quarters from Q4 of 2006 through Q3 of 2007. These calls represent a subset of quarterly earnings calls available on Thomson Reuters StreetEvents for which financial data is available from CRSP, Compustat and I/B/E/S.[17] From this initial sample, we remove observations where the CEO does not speak during the question and answer section and where we are unable to obtain financial statement based predictors (i.e., F-Score and Accounting Risk). Our final sample consists of 1,572 conference call observations. We analyze the CEO's speech during the first 5 minutes of the Q&A portion of the call to obtain the archival analog of the vocal based dissonance metric, *COGDIS*.[18]

4.3 Data and Descriptive Statistics for Archival Data

4.3.1 Misreporting Data

[17] See Mayew and Venkatachalam [2011] for a more detailed discussion of the data collection procedures.

[18] To calibrate the speech of each CEO, we use the opening moments of the CEO speech during the presentation portion of the conference call. Our focus on 5 minutes of CEO speech in the Q&A represents a trade off between the costs of creating audio files for analysis versus more generalizability to, for example, other parts of the conference call like the presentation or other executives such as the CFO. Examining differences in dissonance between the presentation and question and answer session and between the CEO and CFO are fruitful areas of inquiry but beyond the scope of this paper.

We use the Audit Analytics restatements database available via WRDS to identify misreporting firms. We query this database in January 2011 to identify any adverse restatements (i.e. restatements identified as having adverse financial statement effects relative to the originally reported financial statements) for the fiscal quarters in our sample. We find that 111 of our sample firm quarters had adverse financial restatements, representing 56 unique firms.[19] We define *ADV_RES* as an indicator variable that takes a value of one for firm quarters that pertain to a subsequent adverse restatement and zero otherwise. While *ADV_RES* captures overstatements in general, it does not distinguish between unintentional errors from intentional misstatements, i.e., where management intentionally intervened to cause the misreporting, a condition necessary to feel cognitive dissonance. True managerial intent is, of course, not possible to observe. However, convictions of fraud and investigations by outside parties *ex post* are more likely to identify instances of intentional misreporting (Hennes et al. [2008]).

To differentiate intentional (i.e. irregularities) from unintentional (i.e. errors) restatements, we obtain a database update at the end of February 2011 directly from Audit Analytics. This update provides more recent information regarding our sample of adverse restatements in terms of whether they were fraudulent in nature, and whether a regulatory investigation or class action lawsuit has commenced subsequent to the restatement announcement.[20] Adverse restatement firm quarters meeting any of these criteria are classified as irregularity restatements. We define an indicator variable, *IRREG_RES*, that takes the value of one for irregularity restatements, zero otherwise. Of the 56 identified adverse restatements, 16 are identified as irregularities, relating to 40 firm quarters (see Panel A of Table 3). The set of adverse restatements that do not meet the definition of irregularities we label errors (*ERROR_RES*).

---

[19] The most recent restatement announcement available at the time of our data query was September 17, 2010. The number of identified restatements may not represent the full set of restatements that will eventually materialize. Our query in January 2011 is roughly three years subsequent to the sample conference calls analyzed. On average, among all restatements provided by Audit Analytics, the length of time between the beginning of a restatement period and the actual restatement filing date is 2.4 years. Additionally, three of these sample firms restated the same fiscal period more than one time, and in those cases we retain the effects of the first restatement for our analysis.

[20] Because the Audit Analytics does not update this database in real time, we also manually searched Edgar, Google and the SEC website for evidence of any of these conditions for each of the adverse restatements in our sample. We identified one adverse restatement that met one of these conditions that was not already identified in Audit Analytics.

To examine the reasonableness of our restatement classification between intentional and unintentional, we investigate the cumulative abnormal returns during the three days surrounding the restatement announcement, separately, for the two types of restatements. Hennes et al. [2008] document that restatements classified as irregularities exhibit substantially larger negative market reactions surrounding the announcement compared to other restatements. We observe these relationships in our restatement sample. In Panel A of Table 3, we find that on average the market response to all 56 adverse restatement announcements is a statistically negative 6.81% (p= 0.015). This overall negative market reaction is driven by the irregularity restatements. For the 16 restatements classified as irregularities, the average market response is -26.26%, which is statistically different from zero (p=0.002) whereas the market response for the 40 error restatement announcements is not statistically different from zero (average abnormal return = 0.70%, p=0.387).

4.3.2 Remaining Data

Our remaining variables, including the Dechow et al. [2011] F-Score (*FSCORE*), are based on data from Compustat, CRSP, I/B/E/S, and Execucomp as needed. The Compustat variables are from the Unrestated As-First-Reported Compustat database to avoid data backfilling that occurs in the Compustat Fundamentals database in the event of a restatement. Additional executive demographic information not available in Execucomp is hand collected when necessary. The commercial misstatement predictor accounting risk (*ACCT_RISK*), developed and sold by Audit Integrity, LLC, identifies the risk of financial report misrepresentation due particularly to overstated (understated) revenue and assets (expenses and liabilities). *ACCT_RISK* is based exclusively on financial statement information (Correia [2010]), is available on a quarterly basis, and is a parsimonious summary metric that has been shown to perform as well as or better than other accounting based prediction models in the literature (Price et al. [2011]; Correia [2010]).[21] A drawback of this measure is that we cannot state precisely which accounting metrics are the critical drivers behind its predictive ability. *ACCT_RISK* ranges from 0 to 100, with low risk

[21] We thank Jack Zwingli of Audit Integrity (now Governance Metrics International) for providing the *ACCT_RISK* data for our academic use.

receiving higher *ACCT_RISK* scores. To be consistent with other restatement predictors, we modify *ACCT_RISK* by subtracting it from 100 so that higher values capture a higher likelihood of an adverse restatement.

4.3.3 Descriptive Statistics

Panel B of Table 3 provides descriptive statistics for the archival sample. We find that 7.1% of our sample observations report an adverse restatement. The cognitive dissonance measure, *COGDIS*, has a mean (standard deviation) of 0.179 (0.076), which is comparable to the 0.217 (0.088) reported in Table 1 Panel A for the laboratory setting. The market capitalization of the median firm in our sample is $1.452 billion ($e^{7.281}$), which is substantially larger than the median market capitalization of an average Compustat firm of $212 million in fiscal year 2006. In this respect, our analysis differs from other papers investigating the determinants of misreporting (e.g., Dechow et al. [2011]; Price et al. [2011]), which commonly use all available Compustat data. While our sample firms are larger, Panel C of Table 3 reveals that the proportion of sample firms across industries is similar to the Compustat population, with the exception of slight over (under) representation in pharmaceuticals (insurance/real estate).

In Panel D of Table 3, we observe several important bivariate correlations. First, our financial statement predictors of misreporting, *FSCORE* and *ACCT_RISK*, are positively correlated as expected, indicating that both variables capture a common construct. Both variables are also positively correlated with irregularity restatements (*IRREG_RES*), consistent with prior literature (Dechow et al [2011], Price et al. [2011]). Regarding our variable of interest, we find a positive and significant association between *COGDIS* and *IRREG_RES*. Several variables, however, are also associated with both *COGDIS* and *IRREG_RES*, suggesting the potential for correlated omitted variables. As in Mayew and Venkatachalam [2012], more volatile firms and firms with poor performance are more likely to restate and have executives exhibiting higher dissonance levels. Moreover, *ACCT_RISK* and *COGDIS* are positively correlated, making it difficult to discern whether financial statement and voice-based predictors are incrementally predictive of misreporting. To draw more definitive conclusions and to quantify the predictive ability of vocal dissonance cues for misreporting, we turn to multiple regressions next.

4.4 Multiple Logistic Regression Results

In Panel A of Table 4 we establish predictive baselines, separately, for each of the voice based and financial statement based predictors of misreporting. In Column A, we begin with a univariate assessment of how well dissonance markers predict adverse restatements (*ADV_RES*). We observe a positive association between adverse restatements and vocal dissonance markers, but the statistical significance is marginal. To assess predictive ability, we use the area under the Receiver Operator Characteristic (ROC) curve, a technique originally used in signal detection theory (see Hosmer and Lemeshow [2000]). ROC curves help assess the overall discriminatory ability of predictor variables as well as facilitate comparison among alternative predictor variables. In simple terms, a ROC curve is a graphical plot of the probability of detecting a true signal (the true positive rate, also referred to as sensitivity) against a false signal (false positive rate, also referred to as 1-specificity). To plot the curve, it is necessary to estimate the sensitivity and specificity for various "cutoff points" used for classifying the continuous predicted probabilities from the logistic regression in a binary fashion. For example, the predicted probabilities from the model estimated in Column A would help identify the sensitivity and specificity for each cutoff point ranging from 0 to 1. The area under the ROC curve (AUC) is a summary of the overall diagnostic accuracy, with values of 0.500 representing chance levels and 1.000 representing a perfectly accurate prediction model. The area under the ROC curve for Column A is 0.552, and marginally rejects chance levels of 0.500 ($p=0.081$).

The weak results are perhaps not surprising because not all adverse restatements are irregularities (Hennes et al. [2008]). In Columns B and C, we re-estimate the model in Column A, but replace the dependent variable *ADV_RES* with an indicator for error (*ERROR_RES*) and irregularity (*IRREG_RES*) restatements, respectively. This dependent variable refinement reveals a positive and statistically significant coefficient on vocal dissonance markers (coefficient = 5.068, $p < 0.05$) *only* for irregularity restatements in Column C, and not statistically significant for error restatements. The classification accuracy is lower (higher) for error (irregularity) restatements in Column B (C) than for all adverse restatements in Column A as evidenced by an AUC of 0.514 (0.610), which is statistically equivalent to (greater than) chance levels. Since managers should feel dissonant only when they have intentionally misreported, observing the predictive ability of vocal dissonance markers *only* for irregularity restatements and *not* for errors is consistent with *COGDIS* capturing the construct of cognitive dissonance.

Next, we consider an alternative variable for irregularities, i.e., the stock price response to the restatement announcement (Hennes et al. [2008]). If intention is a necessary condition for feeling dissonant, and intentional misstatements tend to have larger price drops on announcement than other restatements, we should observe more vocal dissonance for restatements with larger price declines at the announcement date. To test this assertion, we isolate the 111 firm quarter observations that were adversely restated (Column D). We then regress negative one times the three day cumulative abnormal return (-1**RESCAR*) at the restatement announcement date on *COGDIS*. We multiply the dependent variable by negative one so that the predicted sign on *COGDIS* is positive like the other columns. We observe a positive and significant coefficient on the vocal dissonance markers (coefficient = 0.622, $p<0.05$). This buttresses our finding that CEOs exhibiting higher levels of vocal dissonance are more likely to report an irregularity restatement. Despite this small sample, dissonance markers explain 5.7% of the variance in returns at the announcement, which occur on average 420 days subsequent to the conference call from which we extract dissonance markers. A drawback of using *ex post* outcomes to identify irregularities out of the set of adverse restatements in our sample is that a number of firms may end up as irregularity restatements eventually but have yet to be identified. However, this measurement error only biases against finding an association.

To compare and contrast the predictive ability of vocal markers and financial predictors, in Columns E-H we model irregularity restatements and the market reaction to adverse restatements solely as a function of the two financial statement based predictors we consider, *FSCORE* and *ACCT_RISK*. In Columns E and G we find that both *FSCORE* and *ACCT_RISK* are statistically significant predictors of irregularities, consistent with prior research (Dechow et al. [2011]; Price et al. [2011]). The AUC for the *FSCORE* (*ACCT_RISK*) only model in Column E (G) is 0.599 (0.690), which rejects chance at conventional levels. A visual representation of the area under the ROC curve for each of the variables is provided in Figure 2. A comparison (untabulated) of the AUC for the irregularity restatement models in Columns C, E and G reveals no statistical difference between AUC^{COGDIS} and AUC^{FSCORE} (p=0.856), AUC^{FSCORE} and AUC^{ACCT_RISK} (p=0.117), but a weak difference between AUC^{COGDIS} and AUC^{ACCT_RISK} (p = 0.078). This implies that the commercial metric *ACCT_RISK* provides a weakly dominating predictor in our sample.

Using market responses to adverse restatements, we find *ACCT_RISK* but not *FSCORE* is associated with the stock market reaction. Comparing the adjusted R^2 across models, the *ACCT_RISK* model in Column H provides the most explanatory power (11.3%), followed by the *COGDIS* model in Column D (5.7%) and finally the *FSCORE* model in Column F (0.7%). This rank ordering matches the ordering of AUC, suggesting that our inferences are similar whether we use adverse irregularity restatements or stock market response to restatements as a proxy for intentional misreporting.

To assess whether the predictive power of *COGDIS* is incremental to financial statement based predictors and robust to the inclusion of control variables, we estimate multiple regressions. We report the results in Panel B of Table 4 where all the specifications are logistic regressions with *IRREG_RES* as the dependent variable. In Column A-C, we examine whether *COGDIS* provides incremental predictive power to each of the financial statement based predictors individually and collectively. The coefficient on *COGDIS* remains positive and statistically significant in each of the columns, suggesting that vocal cues provide distinct predictive information for restatements relative to financial predictors. In Columns D-F, we include control variables for other dissonance drivers, CEO characteristics and analyst monitoring. In Column D, we observe a positive and significant coefficient on *COGDIS* of 4.510, which is comparable to the magnitude of 5.068 observed in the univariate model reported in Column B of Panel A. In terms of control variables, firms with higher abnormal returns in the period leading up to the conference call, firms with lower unexpected earnings, and smaller firms are more likely to encounter an irregularity restatement. In Column E, we include the control variables and both financial statement based predictors, but exclude vocal dissonance markers. The results reveal coefficients of similar magnitude to those observed in Column C, implying the inclusion of these control variables have little impact on the financial statement based predictors.

In Column F, we include *COGDIS* along with both financial statement predictors and control variables. In the presence of control variables, both vocal dissonance markers and financial statement based predictors remain incrementally important. To compare relative marginal effects of vocal dissonance drivers to financial statement based predictors, we compute the differences in predicted probabilities as *COGDIS*, *FSCORE*, and *ACCT_RISK* change in value from one standard deviation below to one standard deviation above their respective unconditional means. Other variables are held at their unconditional means and the dichotomous *ALLSTAR_QUES* is held to one for computation. The predicted probability (not tabled) roughly doubles from 1.95% to 3.88% for *COGDIS*, which is identical to the increase for *FSCORE*. Both of these marginal effects are smaller than the effects from *ACCT_RISK*, which improves roughly four fold from 1.34% to 5.59%. We are unable to speculate why the commercial *ACCT_RISK* outperforms the other variables because its construction is proprietary. Overall, we conclude that the results in Table 4 support H1 and suggest that vocal dissonance markers help predict misreporting, incrementally to financial statement predictors.

4.5 Robustness Tests

To ensure that our results are not confounded by other CEO characteristics, we considered the effects of CEO overconfidence and CEO linguistic usage. It is conceivable that overconfidence could be associated with both restatements and cognitive dissonance. To ensure overconfidence is not a correlated omitted factor confounding our inferences, in our multivariate specifications in Table 4 we include a proxy for overconfidence following Malmendier, Tate and Yan [2011] and Malmendier and Tate [2005]. In particular, we include an indicator variable, *LONGHOLDER,* which identifies managers who, during their tenure, held in-the-money options for excessively long periods of time. Our results (not tabled) suggest that including this overconfidence proxy does not alter our inferences.

Regarding linguistics, the extant literature suggests that the linguistic content is impacted by manager-specific speaking styles (Davis et al. [2011]). Moreover, recent work by Larcker and

Zakolyukina [2011] documents that the linguistic content of earnings conference calls helps predict restatements. To ensure that our vocal dissonance metric is not simply proxying for effects associated with linguistic usage, we include each of the linguistic variables investigated by Larcker and Zakolyukina [2011] into our multivariate specifications of Table 4. Results (not tabled) do not affect our inferences.

4.6 Do all managers experience dissonance from financial misreporting?

Our analysis thus far considers CEO vocal dissonance markers at quarter *t* to predict revelations of financial misreporting in the future. Three key elements underlie this prediction. First, managers undertook actions to misreport at or prior to quarter *t*. Second, those actions manifest in cognitive dissonance during earnings conference call discussions. Third, such actions are eventually revealed via a restatement. While it is plausible that conference call discussions may trigger cognitive dissonance in CEOs who overstated financials, it is not necessary that all CEOs experience dissonance to the same degree. Hence, in this section, we explore i) CEO and organizational characteristics that may influence managers to experience differential levels of dissonance and ii) whether such characteristics moderate the association between cognitive dissonance and financial misreporting.

First, we examine whether older, overconfident CEOs and CEOs who are monitored more experience dissonance differentially. CEOs who are monitored are likely to feel more dissonant because monitors may invoke dissonance, not necessarily related to misreporting, through their active engagement with the CEO. As before, we use the incidence of questioning by an all-star analyst (*ALLSTAR_QUES*) as our proxy for monitoring. If higher quality analysts are better able to uncover key topics and ask probing questions that may invoke dissonance, the presence of an all-star analyst question (*ALLSTAR_QUES*) would be positively associated with dissonance. On the other hand, managers have some control over the analysts they decide to speak with publicly during a conference call (Mayew [2008]). If managers who are more susceptible to feeling dissonance choose to avoid discussions with all-star analysts, the *lack of* discussion with an all-star would indicate dissonance.

Research has shown that older individuals tend to have better ability to control negative emotions (Scheibe and Blanchard-Fields [2009]). This implies that older CEOs may be less likely to feel dissonant.

To capture this differential dissonance due to CEO's age, we use an indicator variable, *OLDCEO*, which equals one if the CEO's age is greater than the sample median CEO age of 55 years, zero otherwise. Finally, overconfident CEOs, by being able to find ways to rationalize negative behavior, may also experience lower dissonance levels. To proxy for overconfidence we use the indicator variable, *LONGHOLDER,* which identifies managers who, during their tenure, held in-the-money options for excessively long periods of time (see Malmendier, Tate and Yan [2011]).

We estimate the following determinant model of cognitive dissonance:

$$COGDIS = \beta_0 + \beta_1 (-1{*}RESCAR) + \beta_2 ADV\_RES + \beta_3 ALLSTAR\_QUES + \beta_4 OLDCEO + \beta_5 LONGHOLDER + \varepsilon \tag{3}$$

Note that equation (3) is essentially a reverse regression of the model specified in equation (1). This specification allows us to consider various characteristics that influence cognitive dissonance in one regression. For the restatement outcomes that invoke dissonance we consider two variables, *-1*RESCAR* and *ADV_RES.* As described earlier *-1*RESCAR* is negative one times the cumulative abnormal stock return surrounding the adverse restatement announcement and *ADV_RES* is an indicator variable identifying firm quarters with an adverse restatement. Since firms with no restatements have no market reaction, we set the cumulative abnormal return to zero for non-restating firms.[22] We expect the coefficient β_1 to be positive, consistent with larger magnitude misreporting (as measured by abnormal stock returns) eliciting higher levels of cognitive dissonance. We also expect the coefficient β_2 to be positive as restating CEOs are likely to feel more dissonant than CEOs who did not restate. Consistent with the discussion earlier we expect the coefficient on *ALLSTAR_QUES* to be either positive or negative and that for *OLDCEO* and *LONGHOLDER* to be negative.

Results of estimating equation (3) are presented in Column (A) of Table 5. Consistent with previous findings, we observe that CEOs of firms that experience higher negative restatement returns exhibit increasing levels of cognitive dissonance. That is, the coefficient on *-1*RESCAR* is positive and

[22] Note that we do not restrict our analysis to restating firms in order to maximize power. Setting the cumulative abnormal return to zero for non-restating firms is similar to the treatment of missing data. In such instances, it is customary to include an indicator variable for missing data, commonly known as the zero-order regression (Greene [1993]). The inclusion of *ADV_RES* implicitly serves this purpose.

significant (coefficient=0.094; $p<0.01$). The coefficient on *ADV_RES* is positive but insignificant, which is not surprising given the inclusion of the more powerful magnitude-of-financial-misreporting proxy, *-1*RESCAR*. Beyond the magnitude of misreporting, we find CEOs who are monitored vial all-star analyst questioning during the conference call exhibit less dissonance (coefficient on *ALLSTAR_QUES* = -0.011, $p <0.05$), as do overconfident CEOs (coefficient on *LONGHOLDER* = -0.014, $p < 0.05$).

Next, we investigate whether these three characteristics moderate the association between dissonance and financial misreporting. Specifically, we interact *-1*RESCAR* with *ALLSTAR_QUES, OLDCEO*, and *LONGHOLDER*, separately and include the interaction terms in equation (3). Consistent with the prediction for the main effects, we predict that the coefficient on the interaction terms will be ambiguous for the *ALLSTAR_QUES* interaction term whereas it will be negative for the *OLDCEO* and *LONGHOLDER* interaction terms. That is, we expect a muted association between dissonance and misreporting when CEOs are older and are overconfident.

Results of this modified specification are presented in column (B) of Table 5. We observe that the interaction terms, except for interaction with the overconfidence proxy, are not statistically significant. For overconfident CEOs, we find that the association between misreporting and dissonance is significantly lower (coefficient on *LONGHOLDER* = -0.426; $p<0.05$). In fact, the coefficient representing the association between financial misreporting and cognitive dissonance, turns negative (0.087 - 0.426 = -0.339) for overconfident CEOs. That is, for overconfident CEOs, the more financial misreporting the less the dissonance. However, because only 12% of our sample observations represent overconfident CEOs, and none of them was involved with an irregularity restatement, we hesitate to over-interpret this finding and caution the reader that this analysis should be viewed as a preliminary exploration.

5. Conclusions

We empirically document that vocal dissonance markers are useful for identifying financial misreporting. In a laboratory setting, we generate a speech sample of misreporters and truth-tellers to provide construct validity for the vocal dissonance marker produced by the LVA software. We find vocal dissonance markers from the early part of the speech samples – the precise time when dissonance should be most pronounced – are positively associated with four measures of dissonance from misreporting. This lends support for the LVA-based cognitive dissonance measure. In an archival setting, we find that cognitive dissonance in CEO speech can predict whether a firm's quarterly financial reports will be adversely restated at better than chance levels. The predictive ability of vocal dissonance markers is incremental to accounting based predictors of adverse irregularity restatements.

This paper provides some of the first archival evidence to suggest that important nonverbal clues to detect financial misreporting are present in earnings conference calls. These results should be informative to investors, analysts and auditors who attempt to use earnings conference calls as an information source for assessing the risk of misreporting. However, we caution the reader of the following limitations. First, while we attribute our findings to cognitive dissonance felt by subjects, it is possible that some unknown emotional factor(s) correlated with this construct accounts for our results. Second, LVA is an emerging technology and, as with most commercial products, its inner workings are proprietary. While our laboratory results suggest the LVA dissonance metrics capture aspects of the construct of cognitive dissonance, we are unable to document the mechanisms by which LVA is able to do so. Finally, we only investigate CEO speech. Our laboratory results suggest the possibility that detection of dissonance from misreporting is not specific to CEOs, thus, there may be important information in CFO speech as well. We leave an examination of relative dissonance in CFO speech, or speech from other corporate officials, for future research.

In spite of these limitations, we view our results as a starting point for expanding our understanding of how nonverbal cues play an informational role in capital markets. Numerous questions remain – i) Do emotions stemming from cognitive dissonance matter more or less than other emotions associated with deception?, ii) Are there important interactive effects between vocal and linguistic deception markers?, and iii) To what extent can humans detect dissonance markers on their own independent of or in conjunction with vocal markers? We view these issues as important areas for future inquiry.

Appendix 1

Variable Definitions – Laboratory Setting

CONFESS	Indicator variable that equals 1 if the participant admitted to misreporting score in debriefing, 0 otherwise.
SCORE	Self reported score as stated in response to the first interview question.
SURVEY	Number of points scored during self-timed 5 minute SAT questions online. Four points are given for every correct answer, -1 point for every incorrect answer, and 0 points for every skipped answer.
USCORE	Unexpected score calculated as *SCORE - SURVEY*
BELPRE	Belief regarding number of correctly answered SAT questions that would be obtained in an additional 5 minute SAT test reported after online SAT question timed exam, but before laboratory portion.
BELPOST	Belief regarding number of correctly answered SAT questions that would be obtained in an additional 5 minute SAT test reported during debriefing after laboratory portion.
BELREV	*BELPOST – BELPRE*
COGDIS, E_COGDIS, L_COGDIS	LVA Ex-Sense Pro-R voice based measure of cognitive dissonance, measured as the number of voice segments registering greater than 120 on the Cognition Level measure during the entire interview session, divided by the total number of voice segments in the total interview session. Voice segments are approximate 2-second voice wave intervals. The prefix *E_(L_)* represents measurement during the first (second) half of the interview.
DISFACTOR	First principal component of *BELREV*, *CONFESS*, and *USCORE*, which generates only a single eigenvalue greater than 1.0.
TIME_MIN	Then length of time in minutes of the entire vocal wave file used to generate *COGDIS.*
SCHOOL	Level of education, that equals 2 if subject is a Sophomore, 3 if a Junior, and 4 if a Senior.
MATH	Number of Math courses taken by participant.
ENGLISH	Number of English courses taken by participant.
AGE	Age of participant in years
FEMALE	Indicator variable that equals 1 if the participant was female, zero otherwise.

Variable Definitions – Archival Setting

ADV_RES	Indicator variable that equals 1 if the firm's quarterly financial statements for the firm quarters in our sample where the related conference call audio was analyzed were restated (i.e. the fiscal quarter end falls between *RES_BEGIN_DATE* and *RES_END_DATE* on Audit Analytics via WRDS), the restatement had an adverse impact on the financial statements (*RES_ADVERSE = 1* on Audit Analytics via WRDS). At the time of our data extraction from WRDS in January of 2011, the most recent restatement filing date was September 17, 2010.
IRREG_RES	Indicator variable that equals 1 if the firm's quarterly financial statements for the firm quarters in our sample where the related conference call audio was analyzed were restated (i.e. the fiscal quarter end falls between *RES_BEGIN_DATE* and *RES_END_DATE* on Audit Analytics via WRDS), the restatement had an adverse impact on the financial statements (*RES_ADVERSE = 1* on Audit Analytics via WRDS) and any of the following "irregularity conditions" hold: the restatement was deemed fraudulent (*RES_FRAUD = 1* on Audit Analytics via WRDS*)*, a regulator began an investigation following the restatement (*RES_SEC_INVESTIGATION = 1* on Audit Analytics via WRDS), or a securities class action lawsuit followed the restatement (*DAYS_TO_SECURITIES_CLASS_ACTION* > *0* on Audit Analytics via Audit Analytics online feed). At the time of our data extraction from WRDS in January of 2011, the most recent restatement filing date was September 17, 2010. We obtained an update to the "irregularity conditions" directly from Audit Analytics in February of 2011.
ERROR_RES	Indicator variable for the subset of adverse restatements that are not irregularities, and hence errors. Measured as *ADV_RES – IRREG_RES*
ACCT_RISK	Accounting Risk, defined as the amount of accounting risk the company faces as of the firm's fiscal quarter end. Accounting risk is a financial statement based predictor of the risk that the financial statements are misreported and is provided by the commercial vendor Audit Integrity, LLC. Values range from 0 – 100, with higher values indicating more accounting risk.
FSCORE	Scaled probability of misstatement, estimated as the predicted probability of misstatement scaled by the unconditional probability of misstatement from Dechow et al. [2011] Table 7 Panel A Model 1. The predicted probability is equal to ($e^{\text{predicted\_value}}$ /(1+$e^{\text{predicted\_value}}$)) where the predicted_value = -7.893 + 0.790*RSST Accruals + 2.518*Change in Receivables + 1.191* Change in Inventory + 1.979*% Soft Assets + 0.171*Change in Cash Sales – 0.923*Change in Return on Assets + 1.029 * Actual Issuance. Variable definitions in the prediction equation are quarterly versions of the annual definitions used in Dechow et al. [2011], where changes are derived from the seasonal quarter. The unconditional probability is 494+(494+132,967) = 0.003701. All input variable for calculating predicted_value are winsorized at the 1% and 99% level and come from the Compustat As-First-Reported database.
COGDIS	LVA Ex-Sense Pro-R voice based measure of cognitive dissonance, measured as the number of voice segments registering greater than 120 on the Cognition Level measure from management speech during the quarterly earnings conference call, divided by the total number of voice segments from management speech during the quarterly earnings conference call. Voice segments are approximate 2-second vocal wave intervals.
RET	Current year market adjusted buy and hold stock return as estimated from CRSP, where market adjustment is based on the CRSP value weighted index. Buy and hold return is calculated for the trading days spanning the four fiscal quarters ending at quarter *t* for firm *i*. Variable is winsorized at the 1% and 99% level to mitigate undue outlier effects.
VOL	Stock return volatility, measured as the standard deviation of daily stock returns over the half year period (trading days -127 to ,-2 relative to the conference call date). Variable is winsorized at the 1% and 99% level to mitigate undue outlier effects.
UE	Unexpected earnings at fiscal quarter end, measured as the difference between actual I/B/E/S earnings per share and I/B/E/S analyst summary consensus median earnings per share scaled by price per share two days before the conference call. Variable is winsorized at the 1% and 99% level to mitigate undue outlier effects.
ROA	Return on assets, measured as income before extraordinary items divided by total assets at the beginning of the quarter as provided on the Compustat As-First-Reported database. Variable is

	winsorized at the 1% and 99% level to mitigate undue outlier effects.
AGE	Age of the CEO in years as of fiscal quarter end, as identified by Execucomp or hand collected as necessary.
OLDCEO	Indicator variable that equals one if the *AGE* is greater than the sample median of 55 years, and zero otherwise.
TENURE	Number of years the CEO has been employed by the firm as of the fiscal quarter end, as identified by Execucomp or hand collected as necessary.
RESCAR	Three day cumulative abnormal return centered on the restatement announcement date (which is the variable *FILE_DATE* in Audit Analytics), where the CRPS value weighted index serves as the expected return. Only defined for firm quarters where an adverse restatement was reported.
LONGHOLDER	Proxy for overconfidence based on the holding of in-the-money options following Malmendier et al. [2011] and Malmendier and Tate [2005]. *LONGHOLDER* equals 1 if the CEO, at any point during their tenure at the firm, holds an option until the year of expiration even though the option is at least 40% in the money entering the final year of the option, and zero otherwise.
ALLSTAR_QUES	Proxy for high quality analyst monitoring, measured as an indicator that equals 1 if an Institutional Investor all-star analyst asked a question during the earnings conference call and zero otherwise. All-star analysts are identified by whether the analyst was ranked as a top analyst in either the October 2006 or October 2007 issue of Institutional Investor magazine as our sample period is during calendar year 2007. Names of these all-stars are then compared with the names of the analysts asking questions on each of the conference call transcripts.
BM	The ratio of the book value of equity to the market value of equity at fiscal quarter end.

References

BARTH, M. E., W. H. BEAVER, J. R. M. HAND, and W. R. LANDSMAN. 'Accruals, accounting-based valuation models, and the prediction of equity values.' *Journal of Accounting, Auditing and Finance*, 20 (2005): 311-345.

BENEISH, M.D. 'Detecting GAAP violation: Implications for assessing earnings manageement among firms with extreme financial performance.' *Journal of Accounting and Public Policy* 16 (1997): 271-309.

BOERSMA, P, and D. WEENINK. 'Praat: Doing phonetics by computer (version 5.2.05)' [Computer program], 2010. Available at http://www.praat.org.

BOND, C. F., and B. M. DEPAULO. 'Accuracy of deception judgments.' *Personality and Social Psychology Review* 10 (2006): 214-234.

BRAZEL, J., K. JONES and M. ZIMBELMAN. 'Using nonfinancial measures to assess fraud risk.' *Journal of Accounting Research* 47(2009): 1135-1166.

BURNS, M., K. MOFFITT, W. FELIX and J. BURGOON. 'Using lexical bundles to discriminate between fraudulent and non-fraudulent financial reports,' Unpublished paper, University of Arizona, 2010.

COOPER, J. Cognitive Dissonance Fifty Years of a Classic Theory. Los Angeles, CA: Sage Publications, 2007.

CORREIA, M. 'Political connections, SEC enforcement and acounting quality.' Unpublished paper, London Business School, 2010.

DAVIS, A, W. GE, D. MATSUMOTO, and J. ZHANG. 'The effect of managerial style on the tone of earnings conference calls.' Unpublished paper, University of Oregon, University of Washington, and University of British Columbia, 2011.

DAVIS, A., J. PIGER, and L.SEDOR. 'Beyond the numbers: Managers' use of optimistic and pessimistic tone in the earnings press release.' Unpublished Paper, University of Oregon and University of Washington, 2008.

DECHOW, P.M., W. GE, C.R. LARSON, and R.G. SlOAN. 'Predicting material accounting misstatements.' *Contemporary Accounting Research* 28 (2011): 17-82.

DEMERS, E. and VEGA, C. 'Soft information in earnings announcements: News or noise?,' Unpublished paper, INSEAD, 2008.

DEPAULO, B. M., J. J. LINDSAY, B. E. MALONE, L. MUHLENBRUCK, K. CHARLTON, and H. COOPER. 'Cues to deception.' *Psychological Bulletin* 129 (2003): 74-118.

DYCK, A., A. MORSE, and L. ZINGALES. 'Who Blows the Whistle on Corporate Fraud?' *Journal of Finance* 65 (2010): 2213-2253.

ELKINS, A. 'Evaluating the credibility assessment capability of vocal analysis software,' Forty-Third Annual Hawaii International Conference on System Sciences, Koloa, Kauai, Hawaii, 2010. Available at http://www.hicss.hawaii.edu/Reports/FullProceedings.pdf.

ELKINS, A., and J. BURGOON. 'Validating vocal analysis software to assess credibility in interpersonal interaction: A multilevel factor analytic approach,' Unpublished paper, University of Arizona, 2010.

ELLIOT, A. J., and P. G. DEVINE. 'On the motivational nature of cognitive dissonance: dissonance as psychological discomfort,' *Journal of Personality and Social Psychology* 67 (1994): 382-394.

ELLIOTT, B., F. HODGE, and L. SEDOR. 'Using online video to announce a restatement: Influences on investment decisions and the mediating role of trust.' *The Accounting Review* (2011), Forthcoming.

ENGELBERG, J. 'Costly information processing: Evidence from earnings announcements,' Unpublished paper, Northwestern University, 2008.

EVANS, J., R. L. HANNAN, R. KRISHNAN, and D. MOSER. 'Honesty in managerial reporting.' *The Accounting Review* 76 (2001): 537-559.

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): 915-953.

FESTINGER, L. A Theory of Cognitive Dissonance. Stanford, CA: Stanford University Press, 1957.

FESTINGER, L., and J. M. CARLSMITH. 'Cognitive consequences of forced compliance.' *Journal of Abnormal and Social Psychology* 58 (1959): 203-210.

FRIEDMAN, H. Securities regulation in cyberspace, 3rd ed. Aspen Publishers, 2010.

GAMER, M., H. RILL, G. VOSSEL, and H.W.GODERT. 'Psychophysiological and vocal measures in the detection of guilty knowledge.' *International Journal of Psychophysiology* 60 (2006): 76-87.

GRAHAM, R. 'Theory of cognitive dissonance as it pertains to morality.' *Journal of Scientific Psychology* (2007): 20-23.

GREENE, W. H. Econometric analysis. 4th ed. New Jersey: Pearson Education Inc., 1993.

HALES, J., J. HOBSON, and R. RESUTEK. 'The Dark Side of Socially Mediated Rewards: How Social Status and Narcissism Affect Managerial Reporting,' Unpublished Paper, Georgia Tech University, 2011.

HAN, Y., and J. NUNES. 'Please read the signal but don't mention it: How acknowledging identity signals leads to embarrassment.' *Advances in Consumer Research* 37 (2010), eds. Margaret C. Campbell, Jeff Inman, and Rik Pieters, Duluth, MN: Association for Consumer Research.

HARMON-JONES, E. 'Contributions from research on anger and cognitive dissoance to understanding the motivational function of asymmetrical frontal brain activity.' *Biological Psychology* 67 (2004): 51-76.

HARMON-JONES, E., and J. MILLS. Cognitive Dissonance: Progress on a Pivotal Theory in Social Psychology. Washington, DC: American Psychological Association, 1999.

HENNES, K., A. LEONE, and B. MILLER. 'The Importance of Distinguishing Errors from Irregularities in Restatement Research: The Case of Restatements and CEO/CFO Turnover.' *The Accounting Review* 83 (2008): 1487-1519.

HENRY, E. ‘Are investors influenced by how earnings press relesases are written?’ *Journal of Business Communication* 45 (2008): 363-407.

HIRSCHBERG, J. ‘Deceptive speech: Clues from spoken language.’ In F. Chen and K Jokinen (Eds.): *Speech Technology* (2010), 79-88.

HOLLANDER, S., M. PRONK, and E. ROELOFSEN. ‘Does silence speak? An empirical analysis of disclosure chocies during earnings conference calls.’ *Journal of Accounting Research* 48 (2009): 531-563.

HORVATH, F. S. ‘Effect of different motivational instructions on detection of deception with the psychological stress evaluator and the galvanic skin response.’ *Journal of Applied Psychology* 64 (1979): 323-330.

HOSMER, D., and S. LEMESHOW. Applied logistic regression. New York: Wiley-Interscience Publication, 2000.

HUMPHERYS. S.L, K. MOFFITT, M. BURNS, J. BURGOON, and W. FELIX. ‘Identification of fraudulent financial statements using lingustic credibility analysis.’ *Decision Support Systems* 50 (2011): 585-594.

JAVERS, E. Broker, Trader, Lawer, Spy: The Secret World of Corporate Espionage. New York, NY: HarperCollins Publishers, 2010.

JENSEN, M. T. MESERVY, J. BURGOON, and J. NUNAMAKER. Video-based deception detection. In H. Chen and C.C. Yang (Eds.): Intelligence and Security Informatics Techniques and Applications. Hillsdale, NJ: Springer Berlin/Heidelberg: 233-259, 2008.

JUSLIN, P. N., and P. LAUKKA. ‘Communication of emotions in vocal expression and music performance: Different channels, same code?’ *Psychological Bulletin* 129 (2003): 770-814.

JUSLIN, P. N., and K. R. SCHERER. Vocal expression of affect. In The new handbook of methods in nonverbal behavior research, edited by J. A. Harrigan, R. Rosenthal and K. R. Scherer. Oxford: Oxford Press, 65–135, 2005.

KOTHARI, S. P., X. LI, and J. SHORT. ‘The effect of disclosures by management, analysts, and financial press on cost of capital, return volatility, and analyst forecasts: A Study Using Content Analysis.’ *The Accounting Review* 84 (2009): 1639-1674.

LACERDA, F. LVA technology - The illusion of lie detection. In P. Branderud & H. Traunmuller (Eds.), Fonetik 2009. Stockholm: Department of Linguistics, Stockholm University, 220-225, 2009.

LARCKER, D., and A. ZAKOLYUKINA. ‘Detecting deceptive discussion in conference calls,’ Unpublished Paper, Stanford University, 2011.

LI, F. ‘Annual report readability, current earnings and earnings persistence.’ *Journal of Accounting and Economic.* 45 (2008): 221-247.

LI, F. ‘The determinants and information content of the forward-looking statements in corporate filings - A naïve Bayesian machine learning approach.’ *Journal of Accounting Research* 48 (2010): 1049-1102.

LOUGHRAN, T., and B. MCDONALD. . 'Barron's red flags: Do they actually work?' *Journal of Behavioral Finance* 12 (2011a): 90-97.

LOUGHRAN, T., and B. MCDONALD. 'When is a liability not a liability? Textual analysis, dictionaries and 10-Ks.' *Journal of Finance* 66 (2011b): 35-65.

MALMENDIER, U., TATE, G., and YAN, J. 'Overconfidence and Early-life Experiences: The Effect of Managerial Traits on Corporate Financial Policies.' *Journal of Finance* (2011), Forthcoming.

MALMENDIER, U., and G. TATE. 'CEO overconfidence and corporate investment.' *Journal of Finance* 60 (2005): 2660-2700.

MATSUMOTO, D.A., M. PRONK, and E. ROELOFSEN. 'Managerial disclosure vs. analyst inquiry: An empirical investigation of the presentation and discussion portions of earnings-related conference calls.' *The Accounting Review* (2011), Forthcoming.

MAYEW, W.J. 'Evidence of management discrimination among analysts during earnings conference calls.' *Journal of Accounting Research* 46 (2008): 627-659.

MAYEW, W. J., and M. VENKATACHALAM. 'The power of voice: Managerial affective states and future firm performance.' *The Journal of Finance* (2012) , Forthcoming.

MAZAR, N., O. AMIR, and D. ARIELY. 'The dishonesty of honest people: A theory of self-concept maintenance.' *Journal of Marketing Research* 45 (2008): 633-644.

MESERVY, T., M. JENSEN, J. KRUSE, J. BURGOON, and J., NUNAMAKER. 'Automatic extraction of deceptive behaviorl cues from video.' *Lecture Notes in Computer Science* 3495 (2005): 219-227.

MURPHY, P. 'The attitude toward and rationalization of fraudulent financial reporting,' Unpublished paper, Queen's University, 2010.

NEWMAN, M. L., J. W. PENNEBAKER, D. S. BERRY, and J. M. RICHARDS. 'Lying words: Predicting deception from linguistic styles.' *Personality and Social Psychology Bulletin* 29 (2003): 665-675.

OWREN, M. J. 'GSU Praat tools: Scripts for modifying and analyzing sounds using Praat acoustics software.' *Behavior Research Methods* 40 (2008), 822-829.

PALMATIER, J. J. 'Assessing Credibility: ADVA technology, voice and voice stress analysis.' In R.J. Montgomery and W. J. Majeski. *Corporate Investigations*, Tucson, AZ: Lawyers & Judges Publishing Company, Inc: 37-60, 2005.

PATEL, S., K. SCHERER, E. BJORKNER, and J. SUNDBERG. 'Mapping emotions into acoustic space: the role of voice production.' *Biological Psychology* (2011), Forthcoming. Available at http://dx.doi.org/10.1016/j.biopsycho.2011.02.010.

PRICE, R., N. SHARP, and D. WOOD. 'Detecting and predicting accounting irregularaties: A comparison of commercial and academic risk measures.' *Accounting Horizons* (2011), Forthcoming.

PUBLIC COMPANY ACCOUNTING OVERSIGHT BOARD. Observations on auditors' implementation of pcaob standards relating to auditors' responsibilities with respect to fraud. In PCAOB Release No. 2007-001, 2007.

PUBLIC COMPANY ACCOUNTING OVERSIGHT BOARD. Auditing standards related to the auditor's assessment of and response to risk and related amendments to PCAOB standards. In PCAOB Release No. 2010-004, 2010.

PURDA, L., and D. SKILLICORN. 'Reading between the lines: Detecting fraud from the language of financial reports,' Unpublished paper, Queen's University, 2010.

SALTERIO, S. E., and A. WEBB. 'Honesty in Accounting and Control: A discussion of "the effect of information systomes on honesty in managieral reporting: A behavioral perspective."' *Contemporary Accouning Research* 23 (2006): 919-932.

SCHERER, K. R. 'Vocal Affect Expression - A review and a model for future-research.' *Psychological Bulletin* 99 (1986): 143-165.

SCHEIBE, S., and F. BLANCHARD-FIELDS. 'Effects of regulating emotions on cognitive performance: what is costly for young adults is not so costly for older adults.' *Psychology and Aging* 24 (2009): 217-233.

SCHULLER, B., A. BATLINER, S. STEIDL, and D. SEPPI. 'Recognising realistic emotions and affect in speech: State of the art and lessons learnt from the first challenge.' *Speech Communication* (2011), Forthcoming.

SCHULLER, B. 'On the acoustics of emotion in speech: Desperately seeking a standard.' *Journal of Acoustical Society of America* 127 (2010), 1995-1995.

STICE, E. 'The similarities between cognitive dissonance and guilt: Confession as a relief of dissonance.' *Current Psychology* 11 (1992): 69-77.

TETLOCK, P. 'Giving content to investor sentiment: The role of media in the stock market.' *Journal of Finance* 62 (2007): 1139-1168.

TETLOCK, P., M. SAAR-TSECHANSKY, and S. MACSKASSY. 'More than words: Quantifying language to measure firm's fundamentals.' *Journal of Finance*, 63 (2008): 1437-1467.

VRIJ, A., K. EDWARD, K. P. ROBERTS, and R. BULL. 'Detecting deceit via analysis of verbal and nonverbal behavior.' *Journal of Nonverbal Behavior* 24 (2000), 239-264.

VRIJ, A. Detecting Lies and Deceit: Pitfalls and Opportunities. Chichester, UK, Wiley, 2008.

YU, F. 'Analyst coverage and earnings management.' *Journal of Financial Economics* 88 (2008): 245-271.

ZHOU X.H., N.A. Obuchowski, D.M. Obuchowski. Statistical Methods in Diagnostic Medicine. New York: John Wiley and Sons, 2002.

ZUCKERMAN, M., B. M. DEPAULO, and R. ROSENTHAL. Verbal and nonverbal communication of deception. In Advances in experimental social psychology, edited by L. Berkowitz. New York: Academic Press, 1-59, 1981.

ZUCKERMAN, M., and R. DRIVER. Telling lies: Verbal and nonverbal communication of deception. In A. W. Siegman and S. Feldstein (Eds.). *Multichannel Integrations of Nonverbal Behavior*, Hillsdale, NJ: Lawrence Erlbaum: 129-14, 1985.

Figure 1
Timeline of Events for Generation of Laboratory Data

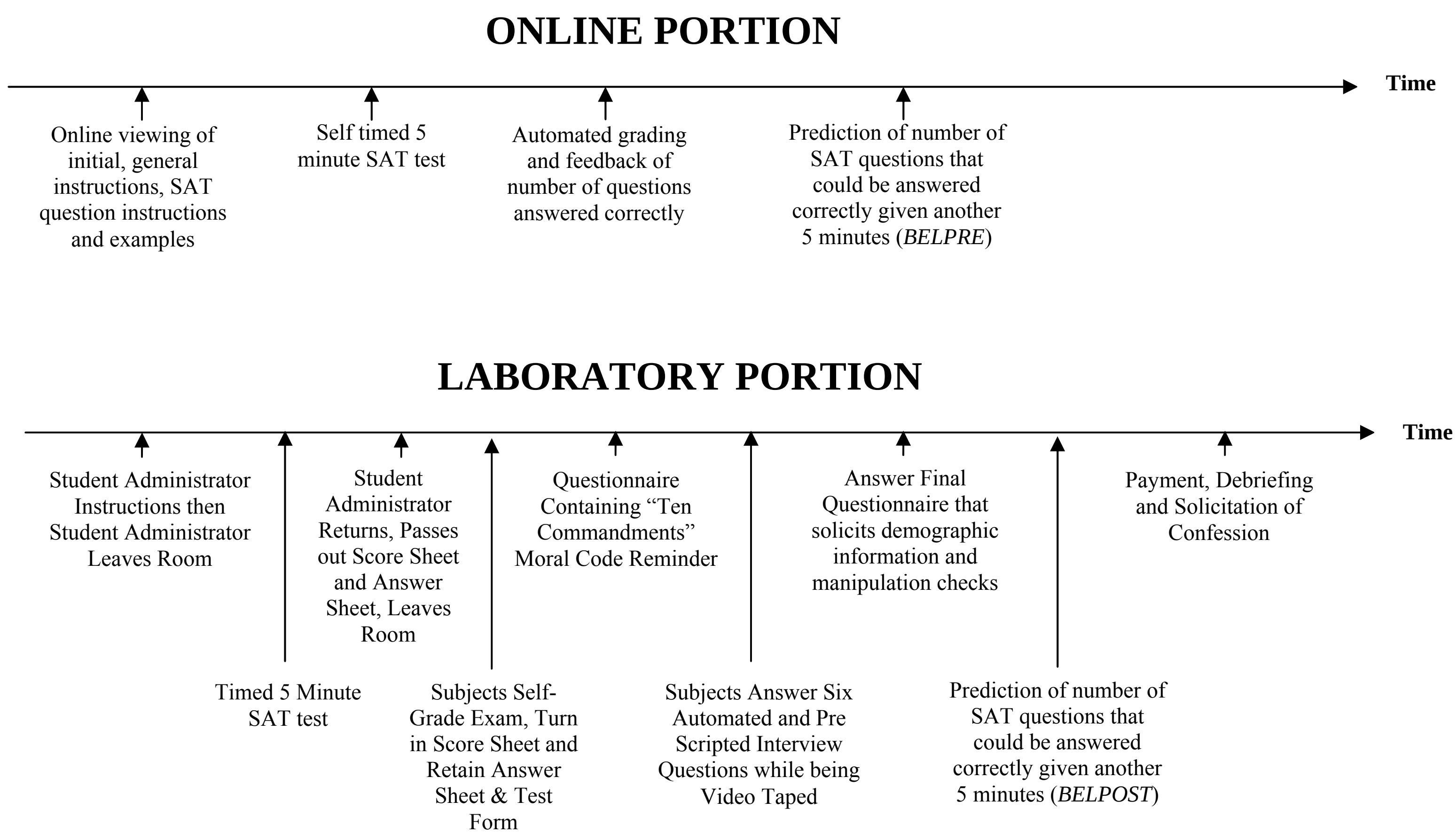

Figure 2
Comparison of Area under the Curve (AUC) for *FSCORE*, *ACCT_RISK* and *COGDIS* in predicting Irregularity Restatements (*IRREG_RES*)

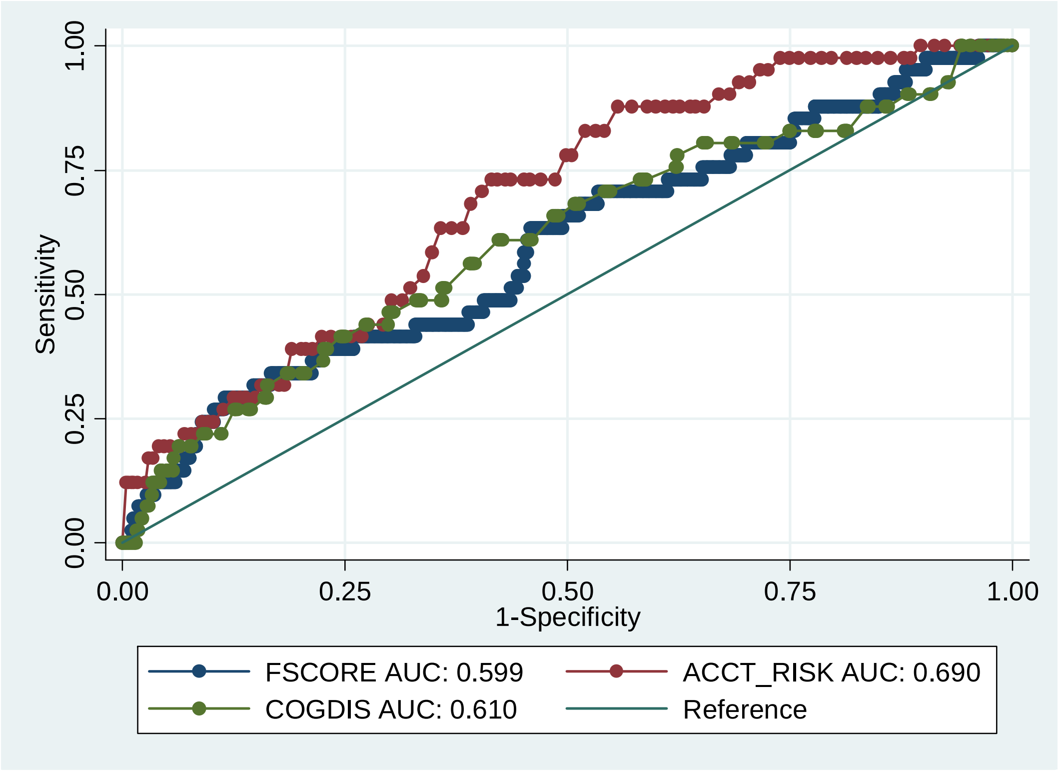

Table 1
Descriptive Statistics and Correlations for Laboratory Generated Data

Panel A: Descriptive Statistics (N=59)[a]

Variable[b]	Mean	Std. Dev	Median	Min	Max
CONFESS	0.322	0.471	0.000	0.000	1.000
SCORE	24.686	11.559	23.000	6.000	70.000
SURVEY	11.627	11.884	8.000	-4.000	61.000
USCORE	13.059	15.967	14.000	-37.000	54.000
BELPRE	6.000	2.205	5.000	2.000	13.000
BELPOST	8.415	3.412	8.000	4.000	22.000
BELREV	2.415	3.467	2.000	-6.000	16.000
DISFACTOR	0.000	1.000	-0.093	-2.701	2.786
COGDIS	0.217	0.088	0.220	0.020	0.484
E_COGDIS	0.214	0.095	0.216	0.000	0.438
L_COGDIS	0.220	0.106	0.200	0.000	0.571
TIME_MIN	2.716	0.791	2.667	1.400	4.733
SCHOOL	2.763	0.916	2.000	1.000	4.000
MATH	2.661	1.469	3.000	0.000	9.000
ENGLISH	1.441	0.992	1.000	0.000	3.500
AGE	19.864	0.937	20.000	18.000	22.000
FEMALE	0.373	0.488	0.000	0.000	1.000

PANEL B: Pearson (Spearman) Correlations above (below) Diagonal Between Proxies for Cognitive Dissonance From Misreporting [c]

	Variable[b]	1	2	3	4	5	6	7
1	***BELREV***		0.128	**0.587**	**0.869**	0.192	**0.320**	0.032
2	***CONFESS***	0.079		0.132	**0.354**	0.173	**0.282**	0.035
3	***USCORE***	**0.680**	**0.222**		**0.870**	0.036	0.162	-0.086
4	***DISFACTOR***	**0.834**	**0.339**	**0.932**		0.158	**0.317**	-0.021
5	***COGDIS***	**0.334**	0.165	0.081	0.190		**0.875**	**0.895**
6	***E_COGDIS***	**0.440**	**0.244**	**0.218**	**0.332**	**0.877**		**0.567**
7	***L_COGDIS***	0.098	-0.001	-0.092	-0.045	**0.847**	**0.538**	

Notes: [a]The full sample is 59 observations gathered from two different Universities. [b] Variable definitions are listed in Appendix 1. [c]Coefficients in bold represent statistical significance at < 10% level.

Table 2
Multivariate Associations Between Proxies for Dissonance from Misreporting and Vocal Dissonance Cues using Laboratory Generated Data

Variable[a]	**Predicted Sign**	(A)	(B)	(C)	(D)
		DISFACTOR	*BELREV*	*CONFESS*	*USCORE*
Intercept	(?)	0.486	9.840	-10.917	20.534
		(2.857)	(9.712)	(9.683)	(48.580)
E_COGDIS	(+)	4.351***	14.262***	6.385**	46.873***
		(1.249)	(5.496)	(3.648)	(14.225)
SCHOOL	(?)	0.203	1.085	-0.358	1.959
		(0.208)	(0.869)	(0.733)	(2.132)
MATH	(?)	-0.031	-0.075	-0.094	-0.308
		(0.078)	(0.293)	(0.205)	(1.087)
ENGLISH	(?)	0.015	-0.321	0.147	1.483
		(0.124)	(0.499)	(0.359)	(1.536)
AGE	(?)	-0.074	-0.609	0.470	-0.648
		(0.155)	(0.522)	(0.564)	(2.605)
FEMALE	(?)	0.117	0.742	0.275	-0.857
		(0.232)	(0.883)	(0.649)	(3.152)
SURVEY	(?)	-0.042***	-0.086**	0.027	-0.950***
		(0.009)	(0.041)	(0.023)	(0.086)
Model Type		OLS	OLS	Logit	OLS
Pseudo R^2 or Adj. R^2		0.379	0.238	0.101	0.552
# of observations		59	59	59	59

Notes: ***, **, * Statistically significant at 1%, 5% and 10% levels in a two-tailed test (one tailed test when predicted). Robust standard errors are presented in parentheses below the coefficient estimates.
[a]Variable definitions are listed in Appendix 1.

TABLE 3
Descriptive Statistics, Correlations and Industry Composition for Archival Data[a]

PANEL A: Descriptive Statistics for Sample Restatements

	Row	# Restatements	Mean *RESCAR*	Median *RESCAR*
Unique adverse restatements identified in Audit Analytics per WRDS affecting sample firm quarters	A	56	-6.81%	-0.49%
Less: Error restatements, which are adverse restatements not identified by Audit Analytics or manual online search as a fraud, as having a regulatory investigation, or being followed by a securities class action lawsuit	B	40	0.70%	0.42%
Equals: Irregularity Restatements	C	16	-26.26%	-17.61%
Tests:				
p-value of t-test Mean *RESCAR* Row A = 0:		0.015		
p-value of t-test Mean *RESCAR* Row B = 0:		0.387		
p-value of t-test Mean *RESCAR* Row C = 0:		0.002		
p-value of t-test Mean *RESCAR* Row B = Mean *RESCAR* Row C:		0.002		

PANEL B: Descriptive Statistics

Variable[b]	N	Mean	Std. Dev	Median	Min	Max
ADV_RES	1,572	0.071	0.256	0.000	0.000	1.000
ERROR_RES	1,572	0.045	0.206	0.000	0.000	1.000
IRREG_RES	1,572	0.026	0.159	0.000	0.000	1.000
RESCAR	111	-0.096	0.216	-0.036	-1.000	0.181
COGDIS	1,572	0.179	0.076	0.172	0.000	0.472
RET	1,572	-0.020	0.383	-0.064	-0.770	1.388
FSCORE	1,572	1.231	0.814	1.091	0.146	4.651
ACCT_RISK	1,572	45.390	26.890	44.000	1.000	100.000
lnMVE	1,572	7.281	1.547	7.169	3.952	11.457
VOL	1,572	0.021	0.009	0.020	0.008	0.051
UE	1,572	-0.001	0.013	0.000	-0.090	0.031
ROA	1,572	0.004	0.044	0.010	-0.206	0.116
AGE	1,572	54.339	7.261	55.000	37.000	83.000
TENURE	1,572	6.087	6.035	5.000	0.000	44.000
ALLSTAR_QUES	1,572	0.309	0.462	0.000	0.000	1.000
BM	1,572	0.444	0.287	0.399	-0.113	1.433

Table 3 (continued)

PANEL C: Industry Composition

	Sample Observations[a]		All Compustat Observations[e]		*ADV_RES* Observations[f]		*IRREG_RES* Observations[f]	
Industry[d]	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
Chemicals	28	1.78	411	1.82	0	0.00	0	0.00
Computers	232	14.76	2,908	12.85	17	15.32	8	19.51
Extractive	55	3.5	904	3.99	0	0.00	0	0.00
Financial	207	13.17	3,050	13.48	12	10.81	6	14.63
Food	21	1.34	401	1.77	2	1.80	0	0.00
Insurance/Real-Estate	115	7.32	2,306	10.19	10	9.01	4	9.76
Manf:ElectricalEqpt	51	3.24	767	3.39	3	2.70	0	0.00
Manf:Instruments	108	6.87	1,062	4.69	9	8.11	7	17.07
Manf:Machinery	28	1.78	544	2.4	5	4.50	4	9.76
Manf:Metal	19	1.21	473	2.09	3	2.70	3	7.32
Manf:Misc.	8	0.51	214	0.95	0	0.00	0	0.00
Manf:Rubber/glass/etc	9	0.57	371	1.64	1	0.90	0	0.00
Manf:TransportEqpt	27	1.72	340	1.5	1	0.90	0	0.00
Mining/Construction	24	1.53	622	2.75	3	2.70	0	0.00
Pharmaceuticals	114	7.25	900	3.98	4	3.60	0	0.00
Retail:Misc.	78	4.96	933	4.12	5	4.50	2	4.88
Retail:Restaurant	17	1.08	286	1.26	0	0.00	0	0.00
Retail:Wholesale	26	1.65	781	3.45	0	0.00	0	0.00
Services	171	10.88	2,064	9.12	13	11.71	0	0.00
Textiles/Print/Publish	79	5.03	845	3.73	12	10.81	7	17.07
Transportation	100	6.36	1,388	6.13	8	7.21	0	0.00
Utilities	49	3.12	658	2.91	3	2.70	0	0.00
Not assigned	6	0.38	405	1.79	0	0.00	0	0.00
Total	1,572	100.00	22,633	100.00	111	100.00	41	100.00

Table 3 (continued)

PANEL D: Pearson (Spearman) Correlations above (below) Diagonal[c]

	Variable[b]	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	***ADV_RES***		**0.783**	**0.594**	N/A	**0.051**	0.026	0.040	**0.069**	**-0.060**	0.027	-0.039	-0.003	0.009	-0.020	**-0.061**	0.002
2	***ERROR_RES***	**0.783**		-0.035	**0.608**	0.014	0.004	-0.004	0.002	-0.019	-0.007	-0.025	0.007	0.020	-0.029	**-0.064**	0.005
3	***IRREG_RES***	**0.594**	-0.035		**-0.608**	**0.065**	0.036	**0.069**	**0.109**	**-0.072**	**0.053**	-0.030	-0.014	-0.011	0.006	-0.014	-0.004
4	***RESCAR***	N/A	**0.615**	**-0.615**		**-0.239**	**-0.260**	-0.083	**-0.342**	**0.190**	**-0.248**	0.061	0.090	**0.342**	-0.026	-0.094	0.127
5	***COGDIS***	**0.046**	0.010	**0.061**	**-0.210**		**-0.052**	-0.018	**0.047**	**-0.114**	**0.093**	**-0.054**	**-0.084**	0.016	-0.004	**-0.069**	-0.015
6	***RET***	0.017	0.004	0.022	**-0.164**	**-0.073**		**-0.081**	**-0.070**	**0.210**	**-0.094**	**0.171**	**0.266**	0.022	-0.035	0.006	**-0.321**
7	***FSCORE***	0.028	-0.008	**0.054**	-0.103	-0.014	-0.041		**0.109**	0.029	-0.014	**-0.166**	**-0.070**	**-0.054**	0.021	-0.032	**0.196**
8	***ACCT_RISK***	**0.067**	0.002	**0.105**	**-0.163**	**0.044**	**-0.054**	**0.052**		**0.043**	0.011	**-0.045**	**-0.088**	**-0.066**	**-0.048**	-0.014	**0.102**
9	***lnMVE***	**-0.060**	-0.015	**-0.076**	**0.193**	**-0.108**	**0.267**	**0.062**	**0.050**		**-0.551**	**0.087**	**0.356**	**0.114**	-0.003	**0.506**	**-0.109**
10	***VOL***	**0.046**	0.001	**0.072**	**-0.245**	**0.055**	**-0.171**	**-0.063**	-0.020	**-0.574**		**-0.192**	**-0.360**	**-0.091**	-0.016	**-0.258**	-0.030
11	***UE***	**-0.063**	-0.028	**-0.064**	0.069	**-0.075**	**0.123**	**-0.081**	-0.010	**0.045**	-0.023		**0.256**	0.020	0.026	0.021	**-0.155**
12	***ROA***	**-0.049**	-0.042	-0.024	0.077	**-0.080**	**0.288**	-0.036	**-0.135**	**0.344**	**-0.193**	**0.195**		**0.048**	0.038	**0.106**	**-0.059**
13	***AGE***	0.008	0.019	-0.012	**0.237**	0.019	0.019	-0.021	**-0.058**	**0.120**	**-0.109**	-0.034	**0.045**		**0.310**	**0.064**	0.041
14	***TENURE***	-0.006	-0.024	0.021	-0.136	-0.026	-0.014	0.017	-0.039	0.013	-0.006	-0.020	**0.042**	**0.221**		-0.021	0.015
15	***ES***	**-0.061**	**-0.064**	-0.014	-0.099	**-0.071**	**0.045**	-0.014	-0.013	**0.505**	**-0.270**	**0.057**	**0.107**	**0.086**	0.021		-0.020
16	***BM***	0.006	0.015	-0.011	0.121	-0.016	**-0.309**	**0.187**	**0.101**	**-0.103**	**-0.092**	**-0.053**	**-0.286**	**0.062**	-0.015	-0.034	

Notes: [a]The number of observations equals 1,647 observations available for voice analysis from Mayew and Venkatachalam [2012] less 13 observations where CEO does not speak, less 52 observations where *ACCT_RISK* is not available, less 10 observations where we cannot calculate *FSCORE* due to missing data in Compustat. [b]Variable definitions are listed in Appendix 1. [c]Coefficients in bold represent statistical significance at < 10% level. [d]Industry definitions follow Barth et al. [2005]. [e]Population is derived from all observations on the annual Compustat database for fiscal year 2006 where SIC code is populated. All continuous variables are winsorized at the 1% and 99% levels to mitigate the effects of outliers. [f]ADV_RES (IRREG_RES) observations pertain to the firm quarters in the sample impacted by the unique adverse (irregularity) restatements identified in Panel A.

Table 4
Logistic Regression Estimation of the Association Between
Proxies for Financial Misreporting, Vocal Dissonance Cues and Financial Statement Based Predictors

PANEL A: Results without including control variables

Variable[b]	Predicted Sign	(A) *ADV_RES*	(B) *ERROR_RES*	(C) *IRREG_RES*	(D) *-1*RESCAR*	(E) *IRREG_RES*	(F) *-1*RESCAR*	(G) *IRREG_RES*	(H) *-1*RESCAR*
Intercept	(?)	-3.055***	-3.221***	-4.601***	-0.024	-4.197***	0.071	-5.044***	-0.042
		(0.332)	(0.398)	(0.558)	(0.050)	(0.388)	(0.052)	(0.657)	(0.048)
Vocal Dissonance Marker									
COGDIS	(+)	2.574*	0.853	5.068**	0.622**				
		(1.607)	(1.982)	(2.464)	(0.331)				
Financial Statement Based Predictor									
FSCORE	(+)					0.418***	0.019		
						(0.173)	(0.030)		
ACCT_RISK	(+)							0.026***	0.003**
								(0.010)	(0.001)
Model Type		Logit	Logit	Logit	OLS	Logit	OLS	Logit	OLS
Pseudo R^2 or Adj R^2		0.005	0.001	0.017	0.057	0.017	0.007	0.049	0.113
# of observations[a]		1,572	1,572	1,572	111	1,572	111	1,572	111
AUC[c]		0.552	0.514	0.610	N/A	0.599	N/A	0.690	N/A
Z-stat for Test AUC = 0.500[c]		1.745	0.379	2.314	N/A	2.086	N/A	5.074	N/A
p-value for Test AUC = 0.500[c]		0.081	0.705	0.021	N/A	0.037	N/A	<0.001	N/A

Table 4 (continued)

PANEL B: Results after including control variables

Variable[b]	Predicted Sign	(A) *IRREG_RES*	(B) *IRREG_RES*	(C) *IRREG_RES*	(D) *IRREG_RES*	(E) *IRREG_RES*	(F) *IRREG_RES*
Intercept	(?)	-5.238***	-5.934***	-6.329***	-1.239	-2.225	-3.029
		(0.641)	(0.983)	(0.984)	(2.572)	(2.606)	(2.643)
Vocal Dissonance Marker							
COGDIS	(+)	5.241**	4.729**	4.959**	4.510**		4.677**
		(2.492)	(2.432)	(2.461)	(2.597)		(2.767)
Financial Statement Based Predictor							
FSCORE	(+)	0.436**		0.324**		0.410**	0.434**
		(0.175)		(0.174)		(0.196)	(0.196)
ACCT_RISK	(+)		0.026***	0.024***		0.028***	0.027***
			(0.010)	(0.010)		(0.009)	(0.009)
Non Misreporting Dissonance Drivers							
RET	(?)				0.807*	0.958**	0.998**
					(0.430)	(0.436)	(0.430)
lnMVE	(?)				-0.435***	-0.568***	-0.563***
					(0.120)	(0.127)	(0.129)
VOL	(?)				-1.110	-1.198	-4.226
					(18.689)	(18.899)	(19.979)
ROA	(?)				2.099	4.237	3.936
					(3.864)	(3.588)	(3.796)
UE	(?)				-10.988*	-8.803	-6.392
					(5.872)	(6.484)	(6.654)
BM	(?)				0.041	-0.336	-0.284
					(1.022)	(0.996)	(1.029)
CEO Characteristics							
AGE	(?)				-0.009	0.006	0.003
					(0.046)	(0.046)	(0.045)
TENURE	(?)				0.011	0.012	0.013
					(0.036)	(0.039)	(0.039)
Monitoring							
ALLSTAR_QUES	(+/-)				0.583	0.846	0.920*
					(0.530)	(0.531)	(0.524)

Table 4 (continued)

Model Type	Logit	Logit	Logit	Logit	Logit	Logit
Pseudo R^2	0.034	0.064	0.075	0.054	0.115	0.128
# of observations[a]	1,572	1,572	1,572	1,572	1,572	1,572
AUC[c]	0.653	0.705	0.723	0.714	0.780	0.787
Z-stat for Test AUC = 0.500[c]	3.196	4.639	5.250	6.495	9.173	8.392
p-value for Test AUC = 0.500[c]	0.001	<0.001	0.002	<0.001	<0.001	<0.001

Notes: ***, **, * Statistically significant at 1%, 5% and 10% levels in a two-tailed test (one tailed test when predicted). Robust standard errors clustered by firm are presented in parentheses below the coefficient estimates. [a]The number of observations equals 1,647 observations available for voice analysis from Mayew and Venkatachalam [2012] less 13 observations where CEO does not speak, less 52 observations where *ACCT_RISK* is not available, less 10 observations where we cannot calculate *FSCORE* due to missing data in Compustat. For tests that examine the stock price response to the restatement announcement, only the 111 adverse restatement observations are utilized. [b]Variable definitions are listed in Appendix 1. [c]The Receiver Operating Characteristic (ROC) curve analysis is used to quantify the accuracy of the logistic prediction equation at classifying participants as having misreported or not. The ROC curve is a graph of the sensitivity versus 1 – specificity of the prediction test. This area measures the global performance of the test. The greater the area under the ROC curve (AUC), the better the performance. The test statistic for testing whether the AUC is statistically different from chance of 0.50 is (AUC – 0.50)/((standard error (AUC)). AUC and standard error (AUC) were obtained from STATA's roctab command. This test statistic is approximately normal (Zhou et al. [2002]), and is therefore reported as a Z statistic, with two sided p-values. All continuous variables are winsorized at the 1% and 99% levels to mitigate the effects of outliers.

Table 5
Effects of CEO characteristics on levels of Cognitive Dissonance

Variable[a]	**Predicted Sign**	(A)		(B)	
		COGDIS		*COGDIS*	
Intercept	(?)	0.186	***	0.185	***
		(0.010)		(0.010)	
-1*RESCAR	(+)	0.094	***	0.087	***
		(0.025)		(0.034)	
ADV_RES	(+)	0.004		0.003	
		(0.010)		(0.010)	
ALLSTAR_QUES	(+/-)	-0.011	**	-0.011	**
		(0.005)		(0.005)	
OLDCEO	(-)	0.005		0.005	
		(0.005)		(0.005)	
LONGHOLDER	(-)	-0.014	**	-0.015	**
		(0.007)		(0.007)	
-1*RESCAR_ALT*ALLSTAR_QUES	(+/-)			0.018	
				(0.037)	
-1*RESCAR_ALT*OLDCEO	(-)			0.036	
				(0.055)	
-1*RESCAR_ALT*LONGHOLDER	(-)			-0.426	**
				(0.172)	
Adjusted R^2		0.016		0.018	
# of observations		1,572		1,572	

Notes: ***, **, * Statistically significant at 1%, 5% and 10% levels in a two-tailed test (one tailed test when predicted). Robust standard errors are presented in parentheses below the coefficient estimates. [a]Variable definitions are listed in Appendix 1. Note that for non-restating firms *-1*RESCAR* is set to zero.