

Budgeting Audit Time: Effects of Audit Step Frame and Verifiability

Eldar M. Maksymov
Arizona State University

Mark W. Nelson
Cornell University

William R. Kinney, Jr.
The University of Texas at Austin

ABSTRACT: Auditors must ensure that their audit plans budget sufficient time for key audit steps. Research has shown that insufficient audit time budgets can be detrimental to audit quality. We examine whether framing audit steps negatively (e.g., assess whether management’s assumptions *are not* appropriate) increases time budgets—particularly for steps in which the auditor perceives that performance quality is less verifiable, and thus most at risk of being performed with low quality. First, we report the results of analyses indicating that, in practice, audit steps are predominantly framed positively, potentially resulting in smaller time budgets. We then report the results of an experiment in which 50 experienced audit managers budget time for an audit program that tests a Level-3 fair-value estimate. Prior research and Public Company Accounting Oversight Board (PCAOB) inspections indicate that this is a challenging audit area, vulnerable to allegations of low audit quality. The results support our predictions and suggest that reframing audit steps negatively would increase audit time budgets—an audit quality indicator—particularly for less-verifiable steps.

Keywords: audit planning; fair-value auditing; frame of audit steps; verifiability of audit steps.

INTRODUCTION

The PCAOB identified the audit time budget as an indicator of audit quality: higher audit time budgets indicate higher audit quality (PCAOB 2015a). More specifically, the PCAOB stated that the amount and allocation of hours to various parts of the audit convey the effort put forth by the auditor, all of which are important inputs into audit quality. This position is supported by prior research (e.g., see Knechel, Krishnan, Pevzner, Shefchik, and Velury [2013] for a review). We propose that two key features fundamental to the auditing context—the framing of audit steps and the perceived verifiability of the performance quality of audit steps—may affect audit budgeting judgments. We examine the effect of audit step frame and verifiability in the setting of auditing fair values, which is an area where step verifiability tends to be low (Bell and Griffin 2012; Christensen, Glover, and Wood 2012; Cannon and Bedard 2017) and auditors are vulnerable to providing insufficient effort (PCAOB 2015b, 2015c, 2015d, 2015e).

Framing refers to alternative yet equivalent ways to describe the same decision task, with alternative frames accentuating positive or negative characteristics of the object of that decision task (Levin, Schneider, and Gaeth 1998). Typically, auditing

We are grateful to Richard C. Hatfield (editor) and the anonymous referees. We also thank Lindsay Andiola, Jeremy Bentley, Rob Bloomfield, Scott Emmett, Devon Erickson, Jackie Hammersley, Max Hewitt, Pat Hopkins, Susan Krische, Christian Leuz, Bob Libby, Dave Ricchiute, Hun Tong Tan, and workshop participants at American University, Brigham Young University, Case Western Reserve University, The George Washington University, Indiana University, Iowa State University, Northwestern University, University of Notre Dame, and the 2016 International Symposium on Auditing Research for comments on this paper. We also thank Jeremy Bentley and Scott Emmett for research assistance.

Funding for the research project described in this article was provided by the Center for Audit Quality. However, the views expressed in this article and its content are those of the authors alone and not those of the Center for Audit Quality.

Editor’s note: Accepted by Richard C. Hatfield.

Submitted: October 2016
Accepted: August 2017
Published Online: September 2017

standards and audit steps are framed positively. For example, the audit is planned and effort is allocated to determine whether a recorded fair value, as well as underlying assumptions, is *reasonable*, rather than *not reasonable* (AS 2502.28-29, PCAOB 2015f; ISA 540.6, IAASB 2008). We manipulate how steps are framed and, based on prior research in psychology, predict that auditors will budget more audit hours when steps are framed negatively. Examining this effect of audit step frame is important: responses from 93 experienced auditors and examination of two Big 4 firms' audit manuals show that, in practice, audit steps are predominantly framed positively, potentially resulting in smaller audit time budgets. We further predict (based on an untested suggestion from the psychology literature) that framing effects are more pronounced for audit steps where auditors believe that the quality of performance is harder to verify. Thus, a negative frame should prompt auditors to budget more hours in order to examine audit steps for which they believe the quality of work will be harder to verify (hereafter, less-verifiable steps).

We test our predictions with an experiment in which 50 audit managers from two Big 4 audit firms plan an audit of an asset classified as Level-3 in the fair-value hierarchy.¹ Their primary task is to determine the number of audit hours necessary for 15 audit steps that, according to their firms' audit manuals, are typically used to audit management's process for determining fair values. We manipulate step frame (positive/negative) between participants and elicit each auditor's perceptions of step verifiability for every one of the 15 steps. We use audit managers as participants because managers tend to be primarily responsible for both planning the audits and reviewing the work of the audit seniors—the latter of which usually perform the audit steps relevant to fair-value audits (Griffith, Hammersley, and Kadous 2015).

As predicted, results indicate a significant main effect of step frame on auditors' planning judgments, with auditors allotting more hours under a negative frame than under a positive frame. Results also provide some support for a significant interaction between frame and perceived verifiability of steps, such that a negative frame increases the planned audit budget more for steps perceived as having lower verifiability (e.g., steps examining management's assumptions).

Our study contributes to the existing audit research literature on framing (e.g., Bedard and Graham 2002; Fukukawa and Mock 2011; Mock and Fukukawa 2016). First, our study extends this literature by showing that the framing of audit steps affects time budgets in the setting of fair-value auditing. We believe that examining the impact of frame in the setting of fair-value auditing is of particular interest to research and practice. Prior research indicates that auditors view fair-value estimates as having a very high risk of misstatement (e.g., see Bell and Griffin 2012; Cannon and Bedard 2017). Thus, when auditing fair values, auditors may assign time budgets that already reflect a very high-risk assessment. This results in time budgets that are already very large and unlikely to increase farther when audit steps are framed negatively.

Second, our study provides some initial evidence that lower verifiability of audit steps may increase this framing effect. More specifically, for steps that reviewers perceive as being more difficult to verify, a negative frame is particularly effective in increasing budgeted audit hours—a measure identified by the PCAOB as one of the key inputs into audit quality (PCAOB 2015a).

The remainder of this paper proceeds as follows: The second section provides background and hypotheses. The third section describes our experimental method. The fourth section provides results and analyses, and the fifth section provides conclusions, implications, and directions for future research.

BACKGROUND AND HYPOTHESES

Fair Values and Audit Planning

Current U.S. and international auditing standards require auditors to determine whether management's estimates are "reasonable" (AS 2502.28-29, PCAOB 2015f; ISA 540.6, IAASB 2008). Fair-value estimates are particularly challenging for auditors because of the high amount of uncertainty present in such estimates (e.g., see Cannon and Bedard 2017; PCAOB 2015b, 2015c, 2015d, 2015e).² In the fair-value hierarchy used by the FASB and IASB, Level-3 estimates involve the greatest amount of subjectivity because they are based on inputs that are not observable (FAS 157 and IFRS 13). Thus, auditing Level-3 fair-value estimates involves significant ranges of potentially acceptable conclusions, providing latitude for psychological factors to affect judgment.

Both U.S. and international standards allow three approaches for auditing fair-value estimates: (1) auditing the management process for determining fair values, (2) developing independent estimates, and (3) auditing subsequent events (AS

¹ The Institutional Review Board at Cornell University, where the data were collected using Qualtrics software, granted approval for this experiment.

² For each of the Big 4, PCAOB inspections identify insufficiency of the extent of testing as one of the deficiencies in testing fair values. For example, the PCAOB mentions "failure to sufficiently test significant assumptions . . . used in developing an estimate" (PCAOB 2015b), failure "to perform any procedures to test valuation of . . . securities, for which fair value was determined by the issuer through an internal valuation model" (PCAOB 2015d), etc.

2502.23, PCAOB 2015f; ISA 540, IAASB 2008). Recent research indicates that auditors predominately choose the approach of auditing management's process (Griffith et al. 2015; Cannon and Bedard 2017).

In addition to identifying what audit steps to perform, auditors budget the amount of time necessary for the subordinates to carry out these steps. The PCAOB specified that an audit's time budget is an input to the audit's quality because the allocation of hours to various parts of the audit conveys the effort the auditor is willing to devote to the step (PCAOB 2015a). The PCAOB (2015a) stated that "the hours that levels of an audit team devote to risk areas can suggest whether audit managers have staffed the audit appropriately to reflect the risk areas identified during the planning phase of the audit, and the extent to which senior members of the team have focused sufficiently on those areas." This position is supported by prior research indicating that higher audit budgets positively affect audit quality (e.g., see Knechel et al. [2013] for a review). For example, research has indicated that time budget pressures may result in trade-offs of audit effectiveness for efficiency (McDaniel 1990) and increase the likelihood of reducing audit quality by acts such as signing off on audit workpapers prematurely, or failing to perform necessary audit steps (e.g., Alderman and Deitrick 1982; Kelley and Margheim 1990; Reckers, Wheeler, and Wong-On-Wing 1997; Asare, Trompeter, and Wright 2000; Agoglia, Hatfield, and Lambert 2015). Thus, time budgets are a reasonable proxy for the audit quality.

Effect of Step Frame on Audit Time Budgeting Judgments

We investigate whether audit time budgeting judgments are affected by the way in which audit steps are presented, or "framed," to the auditor preparing the budget. As stated earlier, *framing* refers to alternative ways of describing the same decision task, with the resulting frames typically differing according to whether they evoke positive or negative connotations (Tversky and Kahneman 1981). Bonner (2007) notes that "attribute framing" is most closely associated with potential framing effects in auditing, where a frame changes only one attribute of a decision task. For example, a positive frame for an audit step is "Assess whether a calculation is accurate." A negative frame for the same step is "Assess whether a calculation is *not* accurate." Importantly, positive and negative frames should be viewed as alternative yet equivalent ways of referring to the same event (e.g., an amount that is not "accurate" is "not accurate"), so framing effects can be unambiguously attributed to psychological processes, as opposed to alternative frames that ask about different events.

Research in psychology suggests that negative wording of a task should encourage auditors to budget more audit hours to that task. Levin et al. (1998) review research in psychology, showing that attribute framing effects occur due to a shift in attention, with people focusing on frame-consistent information and seeking to support the outcome implied by the frame. This effect occurs because framing impacts the encoding and representation of information in memory (Levin and Gaeth 1988). Positive encoding emphasizes positive aspects of information in memory and negative encoding emphasizes negative aspects.³ Thus, psychology research on framing suggests that positive frames may cause auditors to focus on supporting client assertions, while negative frames may cause auditors to complete a more critical and thorough examination of evidence.

Prior research in auditing provides mixed support for this claim and has generally found that framing audit tasks negatively has no effect on auditors' judgment (e.g., Kida 1984; Trotman and Sng 1989; Asare 1992). This is presumably because auditors pay attention to negative information regardless of question framing (Smith and Kida 1991). However, more recent studies suggest that framing may affect auditor judgment under some circumstances. Chung et al. (2013) review research in accounting on the role of biases in auditor judgments and speculate that, in some settings, a positive frame may lead auditors to seek information confirming client assertions and engage in less thorough information processing. At the same time, a negative frame may encourage auditors to avoid seeking information that confirms a client's assertions and to engage in information processing that is more thorough.

Consistent with this suggestion, Martin, Rich, and Wilks (2006) warn that the step wording for auditing complex estimates may encourage auditors to search for evidence that corroborates management's assertions, suggesting a possibility that positively worded audit steps in the setting of auditing complex estimates may lead to under-auditing. However, Fukukawa and Mock (2011) do not find a frame effect on risk assessment in valuation assertion testing. In a follow-up study, Mock and Fukukawa (2015) find a frame effect—but only after including additional information that made the setting less vague. However, the frame manipulation in these studies was achieved by general, rather than attribute, frame manipulation. Frame manipulation in these studies changed more than one attribute, manipulating attention on positive versus negative aspects of the task and manipulating the task's wording to be expressed in noun versus verb form.⁴ Such compound manipulation is

³ Notably, Levin et al. (1998) do not use prospect theory proposed in Tversky and Kahneman (1981; hereafter, TK) to explain the mechanism underlying attribute framing. This is because the risky choice framing discussed in TK involved framing each option in an option set, whereas attribute framing involves manipulating only one attribute of the same choice (see Levin et al. [1998] for further discussion).

⁴ Both studies manipulate frame as follows (emphasis added [as *italic* for form and underline for frame]): Positive frame, "The *valuation* of accounts receivable on the 2004 balance sheet is proper." Negative frame, "The accounts receivable on the 2004 balance sheet is not properly valued."

appropriate for some experiments but precludes the conclusion of which manipulated attribute is responsible for the framing effect.

Bedard and Graham (2002; hereafter, BG) contribute to this research by suggesting that one of the reasons for the prior mixed results on frame effects could be attributed to weak frame manipulations. This explanation is consistent with the suggestion in psychology literature: attribute framing requires a manipulation that is strong enough to affect the encoding and representation of information relevant to the audit step (Levin and Gaeth 1988). In their study, BG use a compound manipulation of frame wherein they vary the wording of individual risk statements, and also manipulate the decision aid's orientation on losses by way of mentioning versus not mentioning the risks of litigation and reputation damage. The results show that, for high-risk clients, auditors identify more risks using negative decision aid orientation over using positive decision aid orientation. We extend this line of research by manipulating only one attribute (positive versus negative frame) in the wording of each audit step of a Level-3 fair-value audit program. This allows us to conclude that positive versus negative frame manipulation *was* responsible for the effect.

In summation, prior research suggests that a negative frame will lead reviewers to expect preparers to process risks related to the audit step more thoroughly and critically, requiring more audit time. Thus, we expect a negative frame to result in a larger audit time budget:

H1: Auditors assign more audit hours when steps are framed negatively than when steps are framed positively.

We believe that examining step framing for the setting of fair-value auditing is an important extension of prior studies on the effect of framing. These prior studies suggest that auditors view the risks of material misstatement in fair values as being very high (e.g., see Christensen et al. 2012; Griffith et al. 2015) and potentially irreducible (e.g., see Bell and Griffin 2012; Cannon and Bedard 2017). Given the auditors' widespread view that the risk of misstatement in fair-value estimates is very high, auditors may tend to develop large time budgets that are not sensitive to frame variations in the audit steps. When a time budget is very high by default, framing audit steps negatively may not increase the time budget further. The settings examined by the earlier studies (e.g., going concern analysis and analysis of accounts receivable) may have been more sensitive to the changes in audit step frames. This is because auditors may not generally view those areas by default as presenting a high risk of material misstatement and requiring high time budgets. Thus, based on the prior research, it is unclear whether we would find support for the framing effect predicted in H1 in the setting of fair-value auditing.

In addition, we believe that examining how a negative frame affects time budgeting is important because audit steps are typically framed positively in practice. We gathered two types of additional data to support this assertion. First, a group of 93 auditors (with an average of six years of public accounting experience) from six large firms respond to a survey question about how audit procedures are usually framed in practice. Auditors selected their response from seven options in counterbalanced order anchored at "Procedures are always framed positively" and "Procedures are always framed negatively." Overall, 91.4 percent of the respondents indicated that procedures tend to be framed positively. Second, two Ph.D. students rated the extent to which audit steps in manuals from two Big 4 firms for auditing fair-value estimates are framed positively versus negatively.⁵ Results indicated that 64 percent of all audit steps are framed (the other 36 percent do not appear to be framed), and that 94 percent of these framed steps are framed positively. These additional analyses support the notion that positively framed steps are the norm—both generally in auditing, and specifically to the guidance used by participants' firms to indicate steps for auditing fair-value estimates.

Interaction between Step Frame and Step Verifiability

We also investigate whether the effect of frame on audit planning judgments differs in a predictable manner, depending on auditors' perceptions of a step's verifiability. Psychology literature suggests that a framing effect requires a range of reasonable choices available to the decision maker, and that this effect has the potential to be more significant when the range of choices is greater (e.g., Levin et al. 1986; Beach, Puto, Heckler, Naylor, and Marble 1996). Auditors may view some steps as more difficult to verify because those steps are more complex—the auditors performing those steps may have had a wider range of audit procedure choices because they had to assess a wider range of risks that were being tested by the audit step. Generally, identifying more specific risks relevant to the step should lead to a greater amount of assigned audit time to allow for obtaining the necessary evidence. Thus, some low-verifiability steps may require a greater amount of audit time. In addition, since a negative frame leads to more thorough processing of the risks, the frame should have a greater effect when audit steps have lower verifiability.

⁵ Kappa coefficients assessing the agreement of the two raters were 0.45 (audit manual of Firm 1; p -value < 0.01) and 0.54 (audit manual of Firm 2; p -value < 0.01), which exceed the threshold of 0.40 generally viewed as indicating moderate reliability (Landis and Koch 1977). Disagreements occurred primarily because one rater initially used the "not framed" category more often than the other rater used it. The raters resolved these disagreements most often by concluding that if in doubt, a procedure should be labeled "not framed."

For example, consider the following audit step: testing the reasonableness of the client's assumption about the level of growth in annual rental revenue. This step is likely difficult to verify because it is complex—the auditor has a wide range of procedure choices to address a wide range of risks. Such risks may include that the management's assumption is based on inappropriate comparisons with other rental properties, or that the current state of the rental market is too unstable to allow a reasonable prediction of annual growth rate. Depending on the identification and assessment of risks, the auditor may choose to undertake only one, or a number of, audit procedures to complete the audit step. The more risks the auditor decides to examine, the more procedures the auditor may have to perform. In this instance, the auditor may choose to examine the relevant clause in the contract, verify the customer's existence and financial standing, review a history of comparable rental agreements, obtain and review local rental market reports, etc. The more procedures the auditor chooses in response to risks, the higher the time budget should be. A negative frame should lead auditors to process risks more thoroughly and critically; so when audit steps have lower verifiability, and thus likely more risks for auditors to consider addressing with procedures, a negative frame should have a greater impact on the time budget.

Conversely, consider the audit step of mathematically recalculating a terminal value of a rental building based on the management's method. This step is likely relatively easy to verify because it is relatively simple—the auditor has a very narrow range of procedure choices to address a very narrow range of risks. There is essentially only one risk that is relevant to this step: that the arithmetic calculation is incorrect. If this calculation was provided in an Excel spreadsheet, then the procedure to address this step would generally involve verifying that the Excel formula uses the right cells and the right arithmetic to calculate the terminal value. Auditors should identify this risk regardless of how thoroughly they are processing risks. Since frame affects thoroughness of risk processing, when audit steps have higher verifiability and, therefore, likely fewer risks for auditors to consider addressing with procedures, a negative frame should have a smaller effect on the time budget.

This combined reasoning suggests that the framing effect's strength may depend on the auditor's perception of the step's verifiability. More specifically, lower verifiability may lead to greater framing effects on budgeted audit time. Therefore, we predict that a negative frame will increase budgeted time more for steps that auditors view as less verifiable:

H2: The extent to which an auditor assigns more audit hours when steps are framed negatively is greater for steps perceived as less verifiable.

METHOD

Overview of Experiment and Participants

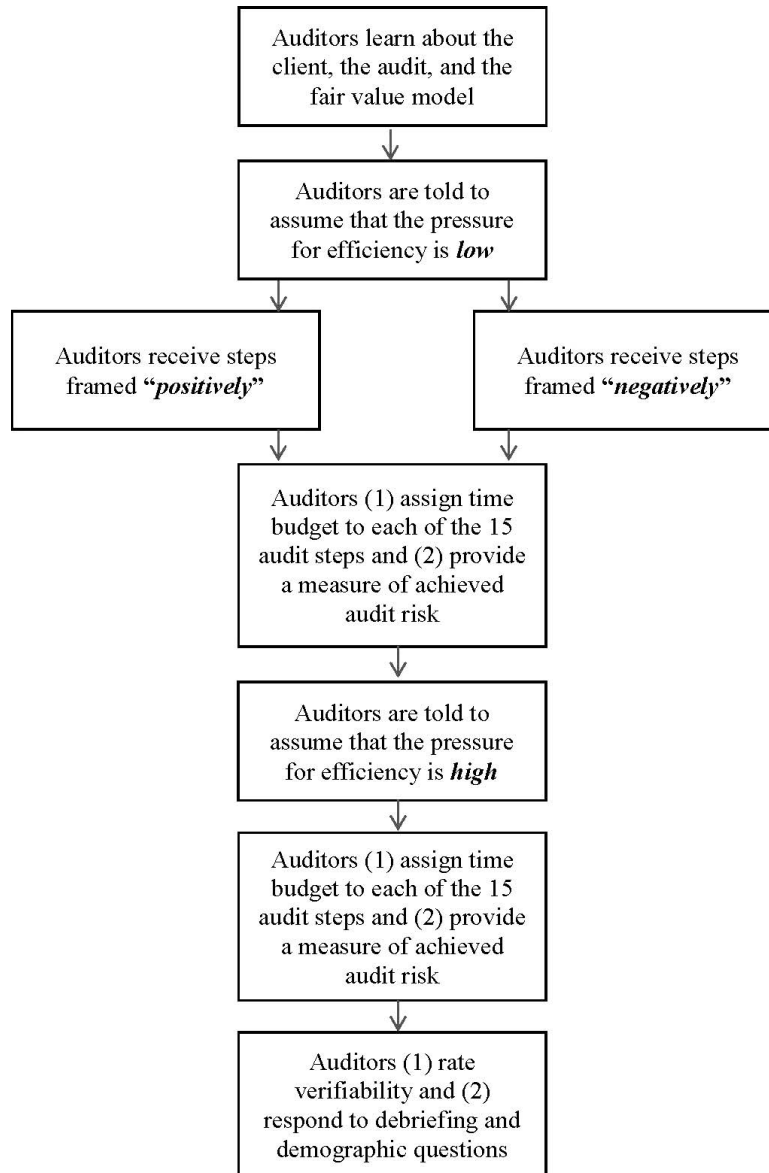
Practicing auditors make audit planning judgments with respect to 15 audit steps in an experiment in which the audit step frame (positive/negative) is manipulated between participants who budget time for each step. Participants budget hours under the assumption that there is relatively low budget pressure (low-efficiency pressure). Then, participants assess the probability that no material misstatement exists (achieved audit assurance), assuming that the work was done within this time budget and no material misstatement was found. Participants then budget for each of the 15 steps again, assuming relatively high pressure to budget few hours (high-efficiency pressure) and, assuming that the work was done within this time budget and no material misstatement was found, assesses the achieved audit assurance again.⁶ Participants then rate the extent to which the audit work quality for each of the 15 steps can be verified in the review process. These verifiability ratings are the basis for our low- or high-verifiability step classification for testing H2. Participants finish the experiment by answering supplemental questions and providing demographic data. Figure 1 outlines our experiment.

The participants include 50 audit managers and senior managers from two Big 4 firms, with an average of 10.7 years of experience.⁷ Participants had worked on an average of 10.3 audits where they audited the valuation model underlying an asset or liability fair value, as well as an average of two audits of real estate investment companies. Participants were recruited by a senior representative of their firm and completed the experiment by accessing the experimental materials online.

⁶ In the low-efficiency pressure, participants are asked to assume a relatively low time pressure. In the high-efficiency pressure, participants are told that the partner asked them to determine the minimum amount of time that could be allocated to each procedure while still providing appropriate assurance that audit objectives have been met. Based on consultations with two audit partners, this order of manipulations matches how auditors might approach the budget—first, consider the work that needs to be done (low-efficiency pressure) and then consider the lowest acceptable amount of time to perform the required work (high-efficiency pressure). Consistent with prior research in accounting (e.g., Houston 1999), efficiency pressure had a significant effect on audit hours: auditors budgeted less time when efficiency pressure was lower (p -value < 0.01). However, we do not find support for our prediction that efficiency pressure will interact with frame. We also find no interaction of efficiency pressure with verifiability. For conciseness, we remove the hypothesis related to efficiency pressure and omit further discussion of the main effect of efficiency pressure.

⁷ One participant did not respond to demographic questions. Results of analyses are similar with that participant excluded, but the main analyses are based on all 50 participants to reflect all data obtained. Twenty-one participants are from one firm, while 29 are from the other. Firm effects are insignificant and do not affect any of our results. Of the 49 participants who responded to demographic questions, 29 are managers and 20 are senior managers. Participants are assigned randomly to experimental manipulations.

FIGURE 1
Experimental Procedure



Dependent Variables

For each of the 15 steps, participants assign a number of planned *audit hours* in 15-minute increments, first under low-efficiency pressure and then under high-efficiency pressure. Measuring the dependent variable in this way allows us to examine our hypotheses at high and low levels of efficiency pressure. We focus our results and perform our tests and analyses using the first measure of audit hours: budgeted hours given low-efficiency pressure. However, we also report whether the results hold under high-efficiency pressure, which is common to audit environments (see [Knechel et al. \[2013\]](#) for a review).

Independent Variables

Step Frame is manipulated between participants by using a positive or negative frame in step wording. An example of a positive framed step is, “Assess whether management’s forecasts and projections have been accurate historically.” An example of the same step under a negative frame is, “Assess whether management’s forecasts and projections have not been accurate

historically.” This frame manipulation is designed to hold constant the participants’ views about the level of assurance provided by the audit, and to produce alternative frames that participants view as complementary and equivalent in terms of the audit task’s implications (e.g., assumptions that are “not reasonable” are not “reasonable,” such that P [assumptions are reasonable] = $1 - P$ [assumptions are not reasonable]).

Perceived Step Verifiability ratings for each of the steps included in Appendix A are measured for each participant. Each participant indicates for each step the extent to which the quality with which the senior performed the step could be verified in the audit review process. A rating of 1 (7) indicates no (complete) ability to verify that the step was performed well.

Task and Step

The task is adapted from training materials used at a Big 4 accounting firm. First, experimental materials were pilot tested with two senior partners specializing in auditing fair values, and were modified based on their comments. Then the materials were pilot tested with 12 audit managers from the two audit firms providing participants, and then again modified. The two audit partners reviewed the final instrument and confirmed that the instrument is realistic and consistent with their firm’s audit approach and policies.

The task invites participants to assume the role of an audit manager responsible for planning the audit of a rental property that is the largest asset on ABC Investment Corporation’s balance sheet. The property is classified as Level-3 in the FASB’s fair-value hierarchy, and participants are provided with the client’s discounted cash flow model used to value the property (calculating a present value of \$10.6 million), and a list of audit steps used by their firm for similar audits.⁸ Participants are informed that comparable properties average 30 to 40 preparer hours (excluding manager and partner review), but vary considerably between clients, inherent and control risks are assessed as sufficiently high to not allow significant modifications to subsequent substantive testing, the partner concluded no specialist was needed for this work, and that the person performing the steps and reporting to the participant would be an audit senior experienced on this audit. Participants are also told that misstatements totaling \$250,000 would be considered material.

After reviewing the client’s schedule of prospective cash flows and the discounted cash flow calculation, participants receive the low-efficiency pressure manipulation and budget audit effort for the 15 audit steps, with the steps framed positively or negatively, depending on the assigned framing condition. All participants receive steps in the same order that is used in the audit firm’s materials from which the case is adapted. In order to preserve the natural order of steps that auditors follow, we do not counterbalance it. We note that while the order could potentially have an effect on the overall level of the assigned audit hours (or on the average verifiability), this impact cannot explain the effects of frame or the interaction of frame and verifiability on audit hours. Participants estimate achieved assurance while assuming that the steps were performed within the time budget and no material misstatement was found. Participants then receive the high-efficiency pressure manipulation, and repeat the process of budgeting audit hours (adjusting the previous hours budgeted under low-time pressure) and estimating achieved audit assurance.

Next, each participant rates the extent to which they believe that the quality with which the senior performed each of the 15 steps could be verified in the audit review process. Data collection concludes with the participants answering some supplemental and demographic questions.

RESULTS

Manipulation Check

To measure whether participants attended to the framing manipulation, a supplemental question was asked to indicate whether audit steps were written to assess whether a particular condition had been met (e.g., “Assess whether management’s forecasts and projections have been accurate historically”), not met (e.g., “Assess whether management’s forecasts and projections have not been accurate historically”), or they did not recall. Forty-two (86 percent) of 50 participants answered the manipulation check correctly, indicating that the frame manipulation was successful. Three participants indicated that they could not remember whether the frame was positive or negative, and one participant did not answer the question. Performance on the manipulation check did not differ between framing conditions (Pearson Chi-squared value = 0.27; $p < 0.60$). Removing the eight participants who failed the manipulation check or did not answer the question does not affect the conclusions. Thus, we base our analyses on all 50 participants.

⁸ We used an asset classified as Level-3 in the FASB’s and IASB’s fair-value hierarchy because we wanted to make sure the audit plan had to address a fair-value estimate that included relatively high uncertainty.

TABLE 1

Descriptive Statistics by Audit Step and by Frame and Repeated Measures ANOVA Examining the Effect of Frame on Hours (Test of H1)

Panel A: Hours Budgeted by Step and by Frame (Hours Mean, [Median], (Standard Deviation))

Audit Step	Positive Frame (n = 24)			Negative Frame (n = 26)			Is Mean Effort Higher under the Negative Frame?
	Mean	Median	SD	Mean	Median	SD	
1	2.60	[2.00]	(1.83)	3.57	[3.50]	(3.17)	Yes
2	0.63	[0.50]	(0.58)	0.88	[0.50]	(0.95)	Yes
3	3.77	[3.50]	(2.87)	5.79	[4.00]	(5.25)	Yes**
4	3.77	[3.00]	(2.58)	3.07	[2.00]	(2.66)	No
5	3.45	[2.00]	(3.29)	3.36	[3.00]	(2.10)	No
6	3.85	[4.00]	(2.26)	4.81	[4.00]	(2.76)	Yes**
7	1.25	[1.00]	(0.90)	1.30	[1.00]	(0.87)	Yes
8	5.54	[5.00]	(2.52)	6.86	[8.00]	(3.60)	Yes*
9	1.72	[2.00]	(1.20)	1.73	[2.00]	(1.23)	Yes
10	1.63	[1.50]	(0.95)	1.74	[2.00]	(0.89)	Yes
11	1.24	[1.00]	(0.91)	1.38	[1.00]	(1.01)	Yes
12	1.15	[1.00]	(0.64)	1.63	[1.25]	(1.52)	Yes*
13	2.35	[2.00]	(1.88)	2.14	[2.00]	(2.05)	No
14	1.29	[1.00]	(1.08)	1.75	[1.00]	(1.24)	Yes*
15	1.29	[1.00]	(0.95)	2.27	[2.00]	(1.28)	Yes***
Average	2.37	[2.00]	(2.36)	2.82	[2.00]	(2.89)	Yes**

*, **, *** Indicate a mixed ANOVA (with participant as the random effect) test significant at the 0.1, 0.05, and 0.01 levels, respectively.

Panel B: Repeated Measures ANOVA Examining the Effect of Frame on Hours (Test of H1)

Source	Numerator DF	Denominator DF	F	p-value
Frame	1	250.073	5.059	< 0.013

Panel A presents descriptive statistics for budgeted hours by frame and step. Auditors assigned hours to 15 audit steps in 0.25 hour increments in the context of auditing the fair value of a real estate investment, first under low-efficiency pressure and then under high-efficiency pressure. The step frame was manipulated between auditors as either positive or negative. Panel B presents the repeated measures ANOVA examining the effect of frame on hours, with step as the repeated measures for each participant.

See Figure 1 for experimental procedure and Appendix A for steps and framing manipulations.

Tests of Hypotheses

Effect of Frame

Table 1, Panel A includes descriptive statistics of auditors' planning judgments by step and step frame. Consistent with H1, when efficiency pressure is low, participants budget significantly more audit hours given a negative frame than they do given a positive frame ($F = 5.059$; $p\text{-value} < 0.013$, Table 1, Panel B).⁹ Under the negative frame, mean effort is higher for 12 of 15 steps (Table 1, Panel A). For the three steps with higher mean effort under the positive frame, the difference is insignificant. At the audit program level, auditors budgeted 31 hours on average for all 15 steps when these steps were framed positively, and 38 hours on average when these steps were framed negatively. Untabulated ANOVA and nonparametric (median, Mann-Whitney, and Kruskal-Wallis) tests show that this difference is significant at a 5 percent level. Untabulated results show that the framing effect is also significant when efficiency pressure is high ($F = 9.377$; $p\text{-value} < 0.01$), and when the two efficiency pressure conditions are combined ($F = 13.105$; $p\text{-value} < 0.01$). These results support H1, indicating that auditors plan more audit hours when steps are framed negatively.

⁹ All models include the *Procedure* variable as repeated measures. Including *Procedure* as the fixed effects variable produces very similar results. All p-values examining directional predictions are one-sided.

TABLE 2
Audit Steps' Verifiability Ratings, Ordered from the Highest to Lowest Mean and Median Value
(Score Mean, [Median], (Standard Deviation))

Step	Verifiability Rating		
	Mean	Median	SD
2	6.76	[7.00]	(0.59)
7	6.46	[7.00]	(1.11)
9	6.16	[7.00]	(1.09)
14	6.18	[6.50]	(0.98)
15	6.18	[6.00]	(0.77)
1	6.12	[6.00]	(0.90)
8	5.32	[6.00]	(1.30)
6	5.20	[5.50]	(1.70)
3	5.26	[5.00]	(1.27)
11	5.22	[5.00]	(1.27)
12	5.02	[5.00]	(1.53)
5	5.00	[5.00]	(1.29)
10	4.86	[5.00]	(1.41)
4	4.76	[5.00]	(1.36)
13	4.58	[4.50]	(1.36)
Average	5.54	[5.70]	(1.20)

This table presents descriptive statistics for verifiability ratings, ordered by step from highest to lowest verifiability. Each participant rated verifiability for each step by indicating the extent to which the quality with which the senior performed the step could be verified in the audit review process, with a rating of 1 (7) indicating no (complete) ability to verify that the step was performed well.

See Figure 1 for the experimental procedure and Appendix A for the steps and framing manipulations.

Combined Effect of Frame and Perceived Verifiability

H2 predicts that the effect of frame on the allocation of audit hours is more severe for steps perceived as less verifiable. Table 2 shows the average verifiability rating for each step. Since H2 predicts an ordinal interaction with a specific pattern (greater effect of frame in lower verifiability), we use planned contrast tests as suggested by [Buckless and Ravenscroft \(1990\)](#) to test this hypothesis. This test requires the independent variables to be categorical. Therefore, we transform the continuous verifiability variable into a categorical variable with two values—high or low verifiability—by comparing the step's average verifiability rating with that of all steps combined.¹⁰

This approach yields classifications that appear intuitive and reasonable. Specifically, our approach classifies Steps 3, 4, 5, 6, 8, 10, 11, 12, and 13 as low verifiability because their mean verifiability is below the overall mean of 5.54 (see Table 2). These steps address the model's assumptions (Steps 3, 4, and 5), ensuring that relevant contractual terms have been incorporated into the model (Step 6), ensuring sufficiency and consistency of evidence (Steps 8, 10, and 11), and making high-level judgments about fair value (Steps 12 and 13). Our approach classifies Steps 1, 2, 7, 9, 14, and 15 as having high verifiability because their mean verifiability is above the overall mean of 5.54. These steps address reviewing the historical accuracy of management's forecasts (Step 1), ensuring accuracy of reconciliation and calculations (Steps 2, 7, and 9), and assessing completeness and conformity of disclosures with GAAP (Steps 14 and 15).

Table 3 shows descriptive statistics by frame and verifiability. Table 3, Panel A shows that, in the low-efficiency pressure condition, the effect of frame was slightly higher in high verifiability compared to low verifiability (0.46 hours versus 0.45 hours), contrary to H2.

As we suggest in our theory for H2, the ability to verify the quality of audit step's performance may reflect the extent to which identifying more risks (by way of the negative frame) would lead to more audit hours. When audit budgets can be set without constraints, it is likely that auditors simply assign hours based on the number of risks they identify for each audit step,

¹⁰ Classifying verifiability based on medians produces very similar results: descriptive statistics and analyses support H2 at the same levels of significance as classifying verifiability based on means. However, classifying based on medians results in significant loss of observations that are equal to medians because these observations do not fall under either high or low verifiability. Thus, we focus on means and do not discuss median-based classifications further.

TABLE 3

Additional Descriptive Statistics—Hours by Frame and by Verifiability (Score Mean, [Median], (Standard Deviation))

Panel A: Low-Efficiency Pressure

Verifiability	Positive Frame (n = 24)			Negative Frame (n = 26)			Diff.
	Mean	Median	SD	Mean	Median	SD	
Low (Steps 3, 4, 5, 6, 8, 10, 11, 12, 13)	2.97	[2.00]	(2.55)	3.42	[2.00]	(3.28)	0.45
High (Steps 1, 2, 7, 9, 14, 15)	1.46	[1.00]	(1.28)	1.92	[1.13]	(1.84)	0.46
Average	2.37	[2.00]	(2.26)	2.82	[2.00]	(2.89)	0.45

Panel B: High-Efficiency Pressure

Verifiability	Positive Frame (n = 24)			Negative Frame (n = 26)			Diff.
	Mean	Median	SD	Mean	Median	SD	
Low (Steps 3, 4, 5, 6, 8, 10, 11, 12, 13)	2.21	[2.00]	(1.97)	2.80	[2.00]	(2.82)	0.59
High (Steps 1, 2, 7, 9, 14, 15)	1.10	[1.00]	(0.95)	1.51	[1.00]	(1.67)	0.41
Average	1.77	[1.00]	(1.73)	2.28	[1.50]	(2.51)	0.51

This table presents descriptive statistics for hours by frame and by verifiability for low- (Panel A) and high-efficiency (Panel B) pressure conditions. See Figure 1 for the experimental procedure and Appendix A for the audit steps and framing manipulations.

while being less sensitive to the effectiveness of increasing audit hours in response to identifying more risks. Thus, verifiability is more likely to moderate the effect of frame when auditors experience efficiency pressure, and are thus more sensitive about whether increasing audit hours in response to more risks will actually be effective. Consistent with this reasoning, in the high-efficiency pressure, Table 3, Panel B shows that the frame effect was greater in the low- compared to high-verifiability conditions (0.59 hours > 0.41 hours), consistent with H2. The comparison between the frame effects is similar when hours are averaged between the low- and high-verifiability conditions (0.52 hours > 0.43 hours).

We test H2 using a contrast test suggested by Guggenmos, Piercey, and Agoglia (2018) for testing whether an effect (i.e., frame) is greater at a different level of a variable (i.e., verifiability) with an incorporated main effect of that variable (i.e., verifiability). This contrast uses the following weights: +1 (low verifiability, positive frame), +3 (low verifiability, negative frame), -2 (high verifiability, positive frame), and -2 (high verifiability, negative frame). This pattern of weights incorporates a greater difference between frames at low verifiability, and a smaller difference at high verifiability, consistent with H2. In addition, this pattern reflects an intuitive prediction that, on average, lower-verifiability steps would require more audit hours. This is in line with our theory for H2 that suggests these steps tend to incorporate more audit work to address a greater number of risks. However, we do not separately hypothesize this main effect prediction because it reflects the nature of Level-3 fair-value auditing, and may or may not be relevant to other auditing areas. Future research could examine whether lower-verifiability steps require more audit hours in other auditing contexts as well. However, this question is beyond the scope of this study.

As recommended by Guggenmos et al. (2018), we perform contrast tests only when descriptive statistics match the hypothesized pattern. At low-efficiency pressure, the descriptive statistics do not match the hypothesized pattern; the effect of frame is somewhat smaller when the verifiability of steps is low compared to when it is high (0.45 < 0.46, Table 3, Panel A). In addition, an ANOVA indicates that the Frame × Verifiability interaction is insignificant ($F = 0.52$; $p = 0.236$, Table 4, Panel A). Thus, H2 is not supported at low-efficiency pressure.

We also analyze whether H2 is supported at high-efficiency pressure. The insights from this analysis are somewhat limited, because in our experiment the auditors assessed the time budget at high-efficiency pressure *after* they assessed it at low-efficiency pressure. Thus, auditors' responses at low-efficiency pressure potentially affected their responses at high-efficiency pressure. However, this analysis can be useful because high-efficiency pressure is the setting in which auditors typically operate (e.g., see Houston 1999; Asare et al 2000). In addition, since the Frame × Verifiability interaction at low-efficiency pressure is insignificant, any carryover effect from budgeting at low-efficiency pressure is unlikely to explain a significant interaction under high-efficiency pressure.

TABLE 4
Tests of H2

Panel A: Mixed Model Test Results at Low-Efficiency Pressure

Source	Numerator DF	Denominator DF	F	p-value
Frame	1	181.544	3.832	0.026
Verifiability	1	720.013	117.837	0.000
Frame × Verifiability	1	720.013	0.520	0.236

Panel B: Contrast Test Results at High-Efficiency Pressure

Comparison	Effect	Estimate	SE	DF	t	p-value
Frame Effect at Low Verifiability > Frame Effect at High Verifiability (Contrast Weights: +1, +3, -2, -2) ^a	0.59 > 0.41	2.397	0.214	745.939	11.207	< 0.001
Frame Effect at Low Verifiability, Negative Frame > Positive Frame (Contrast Weights: -1, +1, 0, 0)	2.21 < 2.80	0.722	0.232	257.331	3.111	< 0.001
Frame Effect at High Verifiability, Negative Frame > Positive Frame (Contrast Weights: 0, 0, -1, +1)	1.10 < 1.51	0.279	0.260	462.172	1.073	< 0.284

^a Residual variance F-test for the contrast [+1, +3, -2, -2] is 0.278; p-value = 0.841. The calculation method is described in the text and footnote 10.

Panel C: Mixed Model Test Results at High-Efficiency Pressure

Source	Numerator DF	Denominator DF	F	p-value
Frame	1	175.106	6.607	0.006
Verifiability	1	713.533	114.933	0.000
Frame × Verifiability	1	713.533	2.155	0.072

This table presents H2 tests for the effects of frame and verifiability on budgeted hours. Panel A presents the results of an ANOVA with frame and verifiability as the predictors and budgeted hours as the response at *low-efficiency* pressure, Panel B presents the results of contrast tests of the interaction pattern and simple effects at *high-efficiency* pressure, and Panel C presents the results of the ANOVA with frame and verifiability as the predictors and budgeted hours as the response at *high-efficiency* pressure.

See Figure 1 for the experimental procedure and Appendix A for the audit steps and framing manipulations.

Table 4, Panel B shows that the contrast is significant ($0.59 > 0.41$, $t = 11.207$; p-value < 0.001). As recommended by Guggenmos et al. (2018), we also calculate a residual between-cells variance test to assess whether the pattern of results tested by this contrast completely explains the observations in our data. The residual variance test is insignificant ($F = 0.278$; p-value = 0.841, Table 4, Panel B), indicating that the data fit the pattern tested by the contrast [+1, +3, -2, -2]. We calculate the residual variance test using the following formula outlined by Rosenthal, Rosnow, and Rubin (2000): $F_{residual} = \frac{F_{between} \times DF_{between} - F_{contrast}}{DF_{residual}}$.¹¹ Consistent with this contrast pattern, Table 4, Panel B also shows that the effect of frame is significant when verifiability is low ($2.21 < 2.80$, $t = 3.111$; p-value < 0.001), but *not* when verifiability is high ($1.10 < 1.51$, $t = 1.073$; p-value < 0.284). ANOVA results in Table 4, Panel C show that the effect of the Frame × Verifiability interaction is marginally significant ($F = 2.155$; p-value < 0.072). Overall, these results provide limited support for H2.¹²

¹¹ Where $F_{between}$ is the F-value of the test examining whether the means of any of the four cells defined by our design differ ($F_{between} = 42.051$), $DF_{between}$ is the numerator degrees of freedom for the $F_{between}$ test ($DF_{between} = 3$), and $DF_{residual}$ is the denominator degrees of freedom for the $F_{residual}$ test ($DF_{residual} = DF_{between} - 1 = 2$). The denominator degrees of freedom are the same for both the $F_{between}$ and $F_{residual}$ tests.

¹² We also assessed whether our theory applies to participants' perceptions of verifiability relative to their own average verifiability rating for all 15 audit steps. For each participant, we classified an audit step as "high" ("low") when the participant's rating of that procedure was higher (lower) than that participant's average rating of all 15 steps. The results using this method support H2 (contrast test p-value < 0.01 and residual variance test p-value > 0.1). The effect of frame is significant only when verifiability is low (p-value < 0.01), but not when verifiability is high (p-value > 0.1). The ANOVA interaction of Frame × Verifiability is also significant (p-value < 0.05).

Additional Analysis

Effect of Frame on the Verifiability Measure

Since verifiability is a measured variable, there is a potential concern that frame may affect verifiability and that this may impact our conclusions. In order to examine this concern, we first test whether frame has a significant effect on verifiability ratings. We find that verifiability is lower when frame is negative (5.44 out of 7.00), compared to when frame is positive (5.65 out of 7.00) and that this difference is marginally significant (p -value < 0.094). This finding provides limited support to the idea that participants may view steps as being less verifiable at negative frames.

However, this effect does not impact our analyses or conclusions. First, we examine whether the effect of frame on verifiability affects our analyses and conclusions for H1, which predicts that a negative frame leads to more audit hours. We included a raw verifiability score into the model with frame as the predictor, audit hours as the response, and steps as the repeated measure. We noted that, consistent with H1, the effect of frame on audit hours (after including verifiability into the model) continues to be significant at the 5 percent level. Thus, any effect of frame on verifiability has no impact on H1.

Second, we examined whether frame's effect on verifiability may impact our analyses and conclusions for H2, which predicts that the frame's effect on audit hours will be greater when verifiability is lower. We removed the effect of frame on verifiability by saving the residuals from the model where frame is the predictor, verifiability raw score is the response, and step is the repeated measure. The residuals from this model represent verifiability stripped of any effect of the frame. We then classify steps into high or low verifiability based on how the mean residual for each step compares to the mean residual for all steps combined. Classifying steps in this way results in exactly the same steps being assigned to a high- or low-verifiability group, when using the raw verifiability score in the main analysis. Thus, any effect of frame on verifiability has no impact on H2. In summation, we conclude that although frame has a marginally significant effect on verifiability, this effect does not impact our analyses or conclusions for H1 and H2.

Achieved Audit Assurance

During the experiment, participants estimated *achieved audit assurance* (AAA) under both low- and high-efficiency pressure. This was accomplished by participants' assessing of the probability that no material misstatement exists within ABC's fair values, under the assumption the audit was performed within their time budget and no material misstatement was found. Average estimated AAA was 68.21 and 76.88 under a positive and negative frame, respectively. A one-way ANOVA indicates no significant main effects of frame ($F = 1.766$; p -value < 0.190) on AAA. Thus, although participants modify their audit programs significantly in response to the frame, their judgments of AAA, assuming that the budgeted audit was completed as planned, are not significantly different. This lack of effect of the frame on AAA, despite a higher time budget, potentially indicates that auditors identify more risks under a negative frame. If auditors believe that under a negative frame they need to address more risks, then higher budgeted time may not lead to an increase in AAA. Future research is needed to examine whether this suggestion is supported.

CONCLUSION

This paper reports results of an experiment in which 50 audit managers, with an average of over ten years of audit experience, plan the audit of an investment classified as Level-3 of the FASB's and IASB's fair-value hierarchy. The audit step frame (positive, negative) is manipulated between participants, and audit step verifiability varies across 15 steps, with each step's verifiability calculated as the mean of all participants' assessments of that step's verifiability. Results indicate that more hours are planned under a negative step frame, particularly with respect to steps that auditors perceive as less verifiable—such as examining management's assumptions (e.g., Step 3), which recent research reports is a challenging area for auditors (see [Bell and Griffin 2012](#); [Griffith et al. 2015](#); [Cannon and Bedard 2017](#)).

Our finding of a main effect of the frame and an interaction between the frame and step verifiability add to prior findings about the effect of the frame in other contexts (e.g., [Kida 1984](#); [Trotman and Sng 1989](#); [Asare 1992](#); [Emby 1994](#); [Emby and Finley 1997](#); [Fukukawa and Mock 2011](#); [Mock and Fukukawa 2016](#)). Most notably, we find that the frame has a larger effect for steps in which auditors believe it is more difficult to verify performance quality. From a practice perspective, our results suggest that framing effects potentially could be used by audit firms, as well as by auditing standard setters, to enhance audit effectiveness. This is similar to decision aids that employ "choice architecture" interventions recommended in other contexts ([Thaler and Sunstein 2008](#)). Firms that understand auditors' tendency to be influenced by a step frame could design their Level-3 fair-value audit programs to frame steps negatively, thus "nudging" auditors to plan more audit hours, which the [PCAOB \(2015a\)](#) identified as a measure of audit quality. Our results indicate that negatively framing audit steps will be particularly effective when under budget pressure in increasing planned audit effort for the less verifiable steps that likely involve significant judgment. Our findings, in combination with [Bedard and Graham \(2002\)](#), who report that a negative frame increases

risk identification, also suggest that the insufficient auditing of fair values reported by the PCAOB may in part be due to incomplete identification of risks during audit planning. Thus, audit costs are likely to increase due to higher audit time budgets, mostly for the areas identified as challenging by PCAOB inspections (see PCAOB 2015b, 2015c, 2015d, 2015e). However, higher audit time budgets may help audit firms address fair-value audit deficiencies identified by the PCAOB in these inspections such as the failure to perform sufficient procedures, omission of key procedures, etc.

The interventions suggested by our results would only be useful if framing effects do not dissipate over time, as auditors become accustomed to a change in the step frame. Extant research in psychology reports that repeated exposure to framing, as well as making participants aware of the framing effect, does not eliminate it—and may even increase it (Levin et al. 1986; Levin, Jasper, and Gaeth 1996; LeBoeuf and Shafir 2003; Chong and Druckman 2007; Druckman, Fein, and Leeper 2012; Baden and Lecheler 2012; Lecheler and de Vreese 2011). This is presumably because repeated consideration increases focus on the aspects highlighted by the frame (Chong and Druckman 2007).

Our research is subject to some additional limitations. First, we use a measured rather than manipulated variable when testing the effect of step verifiability. While our results indicate it is unlikely that carryover or selection biases could account for our evidence of the interactive effect of frame and verifiability on audit planning judgments, we cannot preclude that possibility. Second, although the PCAOB views audit hours as an important input into audit quality, auditors tend to believe that the fair-value risks may not be reducible. Thus, it is unclear whether budgeting additional audit time due to a negative frame would actually improve audit quality. We call for future research to examine this important question. Third, in our experiment, the low-efficiency pressure condition always preceded the high-efficiency pressure condition. We have no reason to expect results to differ under an alternative order, but cannot test that assertion with our data. Finally, our study does not provide insight into how the effects we report would vary with variations in other important factors common in the fair-value audit environment, such as complexity, risk of misstatement, materiality of the account, and others. We call for future research into examining these important questions.

REFERENCES

- Agoglia, C. P., R. C. Hatfield, and T. A. Lambert. 2015. Audit team time reporting: An agency theory perspective. *Accounting, Organizations and Society* 44: 1–14. doi:10.1016/j.aos.2015.03.005
- Alderman, C. W., and J. W. Deitrick. 1982. Auditors' perceptions of time budget pressures and premature sign-offs: A replication and extension. *Auditing: A Journal of Practice & Theory* 1 (2): 54–68.
- Asare, S. K. 1992. The auditor's going-concern decision: Interaction of task variables and the sequential processing of evidence. *The Accounting Review* 67: 379–393.
- Asare, S. K., G. M. Trompeter, and A. M. Wright. 2000. The effect of accountability and time budgets on auditors' testing strategies. *Contemporary Accounting Research* 17 (4): 539–560. doi:10.1506/F1EG-9EJG-DJ0B-JD32
- Baden, C., and S. Lecheler. 2012. Fleeting, fading, or far-reaching? A knowledge-based model of the persistence of framing effects. *Communication Theory* 22 (4): 359–382. doi:10.1111/j.1468-2885.2012.01413.x
- Beach, L. R., C. P. Puto, S. E. Heckler, G. Naylor, and T. A. Marble. 1996. Differential versus unit weighting of violations, framing, and the role of probability in image theory's compatibility test. *Organizational Behavior and Human Decision Processes* 65 (2): 77–82. doi:10.1006/obhd.1996.0007
- Bedard, J. C., and L. E. Graham. 2002. The effects of decision aid orientation on risk factor identification and audit test planning. *Auditing: A Journal of Practice & Theory* 21 (2): 39–56. doi:10.2308/aud.2002.21.2.39
- Bell, T. B., and J. B. Griffin. 2012. Commentary on auditing high-uncertainty fair value estimates. *Auditing: A Journal of Practice & Theory* 31 (1): 147–155. doi:10.2308/ajpt-10172
- Bonner, S. E. 2007. *Judgment and Decision Making in Accounting*. 1st edition. Englewood Cliffs, NJ: Prentice Hall.
- Buckless, F. A., and S. P. Ravenscroft. 1990. Contrast coding: A refinement of ANOVA in behavioral analysis. *The Accounting Review* 65 (4): 933–945.
- Cannon, N. H., and J. C. Bedard. 2017. Auditing challenging fair value measurements: Evidence from the field. *The Accounting Review* 92 (4): 81–114. doi:10.2308/accr-51569
- Chong, D., and J. N. Druckman. 2007. A theory of framing and opinion formation in competitive elite environments. *Journal of Communication* 57 (1): 99–118. doi:10.1111/j.1460-2466.2006.00331.x
- Christensen, B. E., S. M. Glover, and D. A. Wood. 2012. Estimation uncertainty in fair value estimates: Implications for audit assurance. *Auditing: A Journal of Practice & Theory* 31 (1): 127–146. doi:10.2308/ajpt-10191
- Chung, J. O., C. P. Cullinan, M. Frank, J. H. Long, J. Mueller-Phillips, and D. M. O'Reilly. 2013. The auditor's approach to subsequent events: Insights from the academic literature. *Auditing: A Journal of Practice & Theory* 32 (Supplement): 167–207. doi:10.2308/ajpt-50328
- Druckman, J. N., J. Fein, and T. J. Leeper. 2012. A source of bias in public opinion stability. *The American Political Science Review* 106 (2): 430–454. doi:10.1017/S0003055412000123

- Emby, C. 1994. Framing and presentation mode effects in professional judgment: Auditors' internal control judgments and substantive testing decisions. *Auditing: A Journal of Practice & Theory* 13 (Supplement): 102–115.
- Emby, C., and D. Finley. 1997. Debiasing framing effects in auditors' internal control judgments and testing decisions. *Contemporary Accounting Research* 14 (2): 55–77. doi:10.1111/j.1911-3846.1997.tb00527.x
- Fukukawa, H., and T. J. Mock. 2011. Audit risk assessments using belief versus probability. *Auditing: A Journal of Practice & Theory* 30 (1): 75–99. doi:10.2308/aud.2011.30.1.75
- Griffith, E. E., J. S. Hammersley, and K. Kadous. 2015. Audits of complex estimates as verification of management numbers: How institutional pressures shape practice. *Contemporary Accounting Research* 32 (3): 833–863. doi:10.1111/1911-3846.12104
- Guggenmos, R. D., D. M. Piercey, and C. P. Agoglia. 2018. Customer contrast testing: Current trends and a new approach. *The Accounting Review* (forthcoming).
- Houston, R. W. 1999. The effects of fee pressure and client risk on audit seniors' time budget decisions. *Auditing: A Journal of Practice & Theory* 18 (2): 70–86. doi:10.2308/aud.1999.18.2.70
- International Auditing and Assurance Standards Board (IAASB). 2008. *Auditing Accounting Estimates, Including Fair Value Accounting Estimates, and Related Disclosures. ISA 540*. New York, NY: IAASB.
- Kelley, T., and L. Margheim. 1990. The impact of time budget pressure, personality, and leadership variables on dysfunctional auditor behavior. *Auditing: A Journal of Practice & Theory* 9 (2): 21–42.
- Kida, T. 1984. The impact of hypothesis-testing strategies on auditors' use of judgment data. *Journal of Accounting Research* 22 (1): 332–340. doi:10.2307/2490714
- Knechel, W. R., G. V. Krishnan, M. Pevzner, L. B. Shefchik, and U. K. Velury. 2013. Audit quality: Insights from the academic literature. *Auditing: A Journal of Practice & Theory* 32 (Supplement): 385–421. doi:10.2308/ajpt-50350
- Landis, J. R., and G. G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33 (1): 159–174. doi:10.2307/2529310
- LeBoeuf, R. A., and E. Shafir. 2003. Deep thoughts and shallow frames: On the susceptibility to framing effects. *Journal of Behavioral Decision Making* 16 (2): 77–92. doi:10.1002/bdm.433
- Lecheler, S., and C. H. de Vreese. 2011. Getting real: The duration of framing effects. *Journal of Communication* 61 (5): 959–983. doi:10.1111/j.1460-2466.2011.01580.x
- Levin, I. P., and G. J. Gaeth. 1988. Framing of attribute information before and after consuming the product. *The Journal of Consumer Research* 15 (3): 374–378. doi:10.1086/209174
- Levin, I. P., J. D. Jasper, and G. J. Gaeth. 1996. Measuring the effects of framing country-of-origin information: A process tracing approach. *Advances in Consumer Research* 23: 385–389.
- Levin, I. P., R. D. Johnson, P. J. Deldin, L. M. Carstens, L. J. Cressey, and C. R. Davis. 1986. Framing effects in decisions with completely and incompletely described alternatives. *Organizational Behavior and Human Decision Processes* 38 (1): 48–64. doi:10.1016/0749-5978(86)90025-7
- Levin, I. P., S. L. Schneider, and G. J. Gaeth. 1998. All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes* 76 (2): 149–188. doi:10.1006/obhd.1998.2804
- Martin, R. D., J. S. Rich, and T. J. Wilks. 2006. Auditing fair value measurements: A synthesis of relevant research. *Accounting Horizons* 20 (3): 287–303. doi:10.2308/acch.2006.20.3.287
- McDaniel, L. S. 1990. The effects of time pressure and audit program structure on audit performance. *Journal of Accounting Research* 28 (2): 267–285. doi:10.2307/2491150
- Mock, T. J., and H. Fukukawa. 2016. Auditors' risk assessments: The effects of elicitation approach and assertion framing. *Behavioral Research in Accounting* 28 (2): 75–84. doi:10.2308/bria-51326
- Public Company Accounting Oversight Board (PCAOB). 2015a. *Concept Release on Audit Quality Indicators*. Washington, DC: PCAOB.
- Public Company Accounting Oversight Board (PCAOB). 2015b. *Report on 2014 Inspection of Deloitte & Touche LLP*. Washington, DC: PCAOB.
- Public Company Accounting Oversight Board (PCAOB). 2015c. *Report on 2014 Inspection of Ernst & Young LLP*. Washington, DC: PCAOB.
- Public Company Accounting Oversight Board (PCAOB). 2015d. *Report on 2014 Inspection of KPMG LLP*. Washington, DC: PCAOB.
- Public Company Accounting Oversight Board (PCAOB). 2015e. *Report on 2014 Inspection of PricewaterhouseCoopers LLP*. Washington, DC: PCAOB.
- Public Company Accounting Oversight Board (PCAOB). 2015f. *Auditing Fair Value Measurements and Disclosures. AS 2502*. Washington, DC: PCAOB.
- Reckers, P. M., S. W. Wheeler, and B. Wong-On-Wing. 1997. A comparative examination of auditor premature sign-offs using the direct and the randomized response methods. *Auditing: A Journal of Practice & Theory* 16 (1): 69.
- Rosenthal, R., R. L. Rosnow, and D. B. Rubin. 2000. *Contrasts and Effect Sizes in Behavioral Research: A Correlational Approach*. Cambridge, U.K.: Cambridge University Press.
- Smith, J. F., and T. Kida. 1991. Heuristics and biases: Expertise and task realism in auditing. *Psychological Bulletin* 109 (3): 472–489. doi:10.1037/0033-2909.109.3.472

- Thaler, R. H., and C. R. Sunstein. 2008. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New Haven, CT: Yale University Press.
- Trotman, K. T., and J. Sng. 1989. The effect of hypothesis framing, prior expectations and cue diagnosticity on auditors' information choice. *Accounting, Organizations and Society* 14 (5-6): 565–576. doi:10.1016/0361-3682(89)90019-6
- Tversky, A., and D. Kahneman. 1981. The framing of decisions and the psychology of choice. *Science* 211 (4481): 453–458. doi:10.1126/science.7455683

APPENDIX A

Audit Steps and Auditor Agreement on Verifiability Level

Step No.	Audit Step [positive (negative) frame]
1	Assess whether management's forecasts and projections have (not) been accurate historically.
2	Assess whether the fair-value measurement reconciles (does not reconcile) to the financial statements.
3	Review the valuation method and assumptions used to develop the fair-value measures to assess whether they are (not) appropriate considering the nature of the asset being valued (required by AU 328.18).
4	Determine whether all (any) significant assumptions underlying the fair-value measurement have (not) been included in management's calculation of fair value. Significant assumptions are those that are: <ol style="list-style-type: none"> (1) sensitive to variation or uncertainty in amount or nature (for example, assumptions about short-term interest rates are less susceptible to variation than long-term rates); (2) susceptible to misapplication or bias; and (3) a small change in the assumption may result in large changes in the value of the asset or liability being measured.
5	Assess whether the significant assumptions used by management in measuring fair value, taken individually and as a whole, provide (do not provide) a reasonable basis for the fair-value measurements and disclosures (required by AU 328.28).
6	Review the related contractual agreements to assess whether all (there are any) relevant contractual terms have (that have not) been incorporated into the model.
7	Mathematically recalculate the terminal value based on the method employed to assess whether the calculation is (not) accurate.
8	Test the inputs used to develop the fair-value measurements and disclosures to assess whether all (any) such inputs have (do not have) sufficient evidence to support them.
9	Evaluate whether the fair-value measurement has (not) been correctly calculated from such inputs and management's assumptions (AU 328.39).
10	Assess whether the evidence regarding this fair-value amount is (not) consistent with other evidence obtained and evaluated during the audit (required by AU 328.47).
11	Conclude whether audit evidence to reduce the risk of significant misstatement in the financial statements to an appropriately low level is (not both) sufficient and competent (required by AU 328.03 and 328.15).
12	Conclude whether management has (not) appropriately applied the highest and best use concept within FAS 157 (ASC Section 820-10-55).
13	Determine whether the fair-value measurement reflects (does not reflect) the value that would be received or paid, in an orderly transaction between market participants, to sell the asset or transfer the liability at the measurement date.
14	Assess whether Level-3 disclosures required by FAS 157 are (not) complete (ASC Section 820-10-50).
15	Evaluate whether the fair-value measurement and all related disclosures in the financial statements are (not) in conformity with GAAP (AU 328.15).

Appendix A lists the 15 audit steps used in the experimental materials. Audit step numbers correspond to Table 1.
