How the Sleeper Effect Influences Auditors’ Evaluations of Audit Evidence

Jennifer McCallen†
Assistant Professor
University of Georgia
jennifer.mccallen@uga.edu

Christopher P. Agoglia
Richard H. Simpson Professor of Accounting
University of Massachusetts Amherst
cpa22@isenberg.umass.edu

G. Bradley Bennett
Associate Professor
University of Massachusetts Amherst
gbbennett@isenberg.umass.edu

August 2019

† Corresponding Author

We are grateful for helpful comments and feedback from Jess Buchanan, Mary Kate Dodgson, Jeff Hales, Jackie Hammersley, David Huber, Yoon Ju Kang, Nikki Mackenzie, Dave Piercey, Shankar Venkataraman, and Rob Whited, and workshop participants at the 2019 Audit Midyear Conference, Georgia Institute of Technology, Georgia State University, the University of Georgia, and the University of Massachusetts Amherst.
How the Sleeper Effect Influences Auditors’ Evaluations of Audit Evidence

ABSTRACT

Auditing standards require auditors to consider management sources when evaluating evidence. However, standards are silent on how auditors should identify and incorporate relevant source information into evidence evaluation, despite the source’s importance to evidence evaluation. The “sleeper effect of the source,” a psychological phenomenon, occurs when the persuasiveness of a weak message increases over time when delivered by a credible source. Research highlights that credibility is multi-dimensional and, importantly, competence, one dimension of credibility, is context-specific. An expert in one area may lack knowledge in other areas. Using an experiment, we find that when a credible source provides weak evidence, auditors are subject to a sleeper effect. Current audit practices regarding management assessments result in auditors focusing on overall credibility, as opposed to context-specific competence, when evaluating evidence. We propose and test a documentation intervention that aids auditors in incorporating context-specific competence into evidence evaluations and moderates the sleeper effect.

Keywords: evidence evaluation; auditor-client interactions; source credibility; context-specific competence; audit documentation; audit quality.
I. INTRODUCTION

Current auditing standards require auditors to consider the credibility of management during both the planning and fieldwork phases of the audit (PCAOB AS 2110; PCAOB AS 2501; PCAOB AS 2502). Auditors’ assessments of management credibility during the planning phase focus on overall credibility, typically considering factors such as education, tenure within the company and/or industry, and any specialist certifications or qualifications. During fieldwork, auditors should incorporate relevant source information when evaluating audit evidence during control and substantive test work (PCAOB AS 2501, PCAOB AS 2502). Yet, auditing standards are silent on how auditors should incorporate source information into evidence gathering and evaluation during fieldwork.

Absent any guidance on how to evaluate and incorporate source information when evaluating evidence during the fieldwork phase (later in the audit), auditors are likely to rely on and incorporate the general, more holistic assessment of management credibility made earlier during the planning phase of the audit. When the auditor’s holistic assessment is that management is credible, this could result in a recently identified psychological phenomenon known as the “sleeper effect of the source,” which occurs when a credible source’s message increases in persuasiveness over time (Albarraén, Kumkale, and Poyner-Del Vento 2017). If a generally credible source provides weak, unconvincing audit evidence, the recipient of the information (e.g., an auditor) associates the credibility of the source with the evidence. Over time, the persuasiveness of the evidence increases as a result of its association with a credible source, despite the evidence itself being unconvincing.

Source credibility is an important factor in the persuasiveness of a message (Eagly and Chaiken 1984), and auditors consider information obtained from credible sources to be more
diagnostic (Hirst 1994) and more reliable (Anderson et al. 1994). However, if the source of
information is considered credible in general but lacks the requisite knowledge and experience
with the area being audited, it may not be appropriate to rely on the provided evidence. While the
current literature on credibility recognizes its multi-dimensional nature (O’Keefe 2016;
Pornpitakpan 2004), this literature fails to consider one important feature that influences one’s
overall perception of credibility: context. Competence, one dimension of credibility, is context-
specific, and one’s competence depends on the relevant area or information of interest.¹ Thus, an
individual can be credible, and competent, from an overall perspective but may not be competent
in all relevant areas (O’Keefe 2016).

When the audit team assesses management during the planning phase of the audit, the
auditor may conclude that a manager is, in general, highly credible. However, that does not mean
the manager is necessarily competent or knowledgeable about every aspect of the business and
the account-specific information the auditor is testing (i.e., in all contexts).² This distinction
between overall credibility and context-specific competence has been largely unexamined in
psychology and accounting research. Yet, it is an important distinction, especially for auditors,
who are required to consider management’s credibility in both assessing overall audit risk
(PCAOB AS 2110) and in evaluating area-specific audit evidence (PCAOB 1105, PCAOB AS
2301, PCAOB AS 2501). If auditors subsequently fail to identify and explicitly consider context-
specific competencies (e.g., “specific knowledge and experience” within the given audit area, as
current standards advise (PCAOB AS 2501.03) and instead focus on their earlier assessment of

---

¹ Other dimensions of credibility include trustworthiness, likeability, attractiveness, authoritativeness, and
objectivity (Pornpitakpan 2004).

² For example, while the auditor may have concluded the accounting manager is credible and knowledgeable about
financial reporting, he may know very little about inventory production cycles. Although the manager lacks
competence regarding inventory production cycles, he can still be considered credible overall.
overall credibility when evaluating the audit evidence, their susceptibility to a sleeper effect is exacerbated and, in turn, audit quality may be compromised.

In this study, we consider how auditors may be vulnerable to a sleeper effect when they initially make a more general, overall credibility assessment prior to, and separate from, evaluating evidence, as it is typically done in practice. Further, we examine how management’s context-specific competence (or lack thereof) can influence auditor judgments of audit evidence when a generally credible source provides information to the auditor. In addition, we propose a theoretically-based documentation intervention to aid auditors in identifying and incorporating relevant source information into evidence evaluation.

We manipulate three independent variables: source assessment timing (the source assessment is either concurrent with or separate from the evaluation of evidence), context-specific competence (the source is more or less competent in the audit area), and documentation intervention (instructions to consider area-specific competence).\(^3\) In the separate evaluation conditions, participants first perform an evaluation of the client source (i.e., management) and subsequently obtain and evaluate the reasonableness of audit evidence for a subjective audit area (i.e., separate evaluation of management and evidence).\(^4\) The sequential order of these assessments is designed to mirror the temporal relationship common in the current audit environment between credibility assessments of client management (often made during planning) and evidence evaluation (performed during fieldwork). In the concurrent evaluation conditions, participants evaluate both the source and audit evidence concurrently. These conditions represent

\(^3\) We include a fourth independent variable (discussed in detail later) which manipulates the order of the audit tasks to control for order effects.

\(^4\) The source’s credibility is high, overall, and held constant across all experimental conditions, although context-specific competence varies depending on the audit area being evaluated.
auditors’ judgments if they were able to perform management assessments and evidence evaluations jointly in practice.\(^5\)

We find that auditors consider evidence to be more persuasive when management and the evidence are evaluated separately, relative to the concurrent evaluation conditions. That is, auditors are susceptible to a sleeper effect of the source and, as a result, they evaluate audit evidence as more reasonable when evaluation of the evidence is more temporally disconnected from the source credibility assessment. This effect is stronger when a credible source has less context-specific competence in an area. However, we find that the sleeper effect is reduced when auditors are provided with documentation instructions aimed at assisting auditors in efficiently identifying and incorporating relevant context-specific competence information into evidence evaluation. Specifically, when auditors are instructed to include relevant source knowledge into audit evidence documentation, auditors no longer consider evidence as more persuasive and their evidence reasonableness assessments are more in line with auditors’ evidence evaluations made in the concurrent evaluation conditions. Further, auditors document significantly more facts regarding both the evidence and source when the documentation intervention is present.

We also provide insight into auditors’ evidence assessments. The documentation intervention first lowers auditors’ assessments of the source’s context-specific competence and, in turn, reduces auditors’ perceptions about the sufficiency and appropriateness of the audit evidence obtained, which jointly mediate auditors’ evidence reasonableness assessments. This aligns with auditing standards, which charge the auditor with “obtaining sufficient appropriate audit evidence to provide a reasonable basis for his or her opinion” (PCAOB AS 1105.04). Thus,

\(^5\) While it may be potentially beneficial to require auditors to perform management assessments during both planning and fieldwork, it would be highly inefficient. Auditors face substantial budget constraints, especially during year-end fieldwork, and often rely on work performed at interim periods (Lopez and Peters 2011; Low and Tan 2011; Messier, Glover, and Prawitt 2017).
we provide evidence that auditors’ consideration of context-specific information directly influences their evaluations of evidence sufficiency and appropriateness, the measures by which evidence should be evaluated according to auditing standards.

This study contributes to both the auditing and psychology literatures. We provide new evidence on the sleeper effect, finding that it not only influences persuasion but also judgment and decision making. We identify a potential threat to audit quality based on current auditing practices, and we provide audit professionals with a reasonable documentation instruction that helps to counteract any potentially negative implications to audit quality from a sleeper effect.

Additionally, both auditing and psychology research have identified the importance of considering how the source of information influences persuasion (Pornpitakpan 2004; O’Keefe 2016). However, this is the first study we are aware of that explicitly examines context-specific competence. Further, we examine how overall management credibility assessments can interfere in the identification of relevant context-specific competence during evidence evaluation. Given the importance of incorporating source information throughout the audit, especially with subjective audit areas, it is important to investigate the potential unintended consequences of general assessments of credibility on more context-specific audit tasks. Finally, we answer a call for more research on how auditors incorporate various sources of evidence, specifically with regard to subjective audit areas (Griffith, Hammersley, and Kadous 2015).

II. LITERATURE REVIEW AND HYPOTHESES DEVELOPMENT

Background

Consideration of Management Sources: Current Auditing Practices

Auditing standards stress the importance of the auditor having a well-planned audit approach prior to performing the work (PCAOB AS 2101). Thus, the audit team spends a
significant amount of time during the planning phase of the audit gathering information to identify risk of material misstatement and appropriately plan the audit. Auditing Standard 2110 – *Identifying and Assessing Risk of Material Misstatement* – addresses auditors’ responsibilities during audit planning, requiring auditors to assess the qualifications of key members of the management team (hereafter, “management”) as part of the preliminary assessment of risk of material misstatement (PCAOB AS 2110). These assessments of management are important to assessing the risk of error or fraud, as management is responsible for preparing the financial statements (PCAOB AS 2110.04-05). Further, management is a primary source of information needed to evaluate financial statement information. If management sources lack credibility, in general, and competence in the area being evaluated, auditors may not be able to rely on the evidence they provide.

Assessments of management during the planning phase of the audit typically include gathering background information on management personnel, including education, tenure within the industry and/or company, and any certifications or specialist qualifications (PCAOB AS 2110; Messier et al. 2017). Auditors may identify and evaluate detailed information, such as specific competencies, as part of the assessment made during audit planning. However, the assessment is performed with the intention of forming an overall conclusion regarding the credibility of key members of management and related risks to the financial statements. As a result of these assessments, auditors are better able to identify any risk factors associated with

---

6 PCAOB Auditing Standard 1105 – *Audit Evidence* – defines audit evidence as “all the information, whether obtained from audit procedures or other sources, that is used by the auditor in arriving at the conclusions on which the auditor’s opinion is based. Audit evidence consists of both information that supports and corroborates management’s assertions.” The standard also explicitly states that the source is an important factor in determining the reliability of evidence (PCAOB AS 1105.08). While auditors gather a variety of evidence types, in this study, we refer to audit evidence as the information obtained from a management source regarding the audit areas evaluated.
management and determine how to best design the nature and extent of audit procedures used during audit fieldwork to mitigate auditors’ detection risk.

In addition to evaluating management during the planning phase of the audit, the auditor is also required to consider management sources during audit fieldwork and evidence evaluation (PCAOB AS 2301, 2501, 2502). While the planning phase precedes fieldwork, planning tasks are generally not performed immediately prior to fieldwork. Audit planning can precede audit fieldwork by as much as six to eight months (Sheriff 1998; Messier et al. 2017). The temporal aspect of the audit plays an important role in how, and what, information about the source originally evaluated during planning is incorporated into evidence evaluation during fieldwork.

During audit fieldwork, auditing standards emphasize the importance of the source when evaluating subjective audit areas, such as audit estimates. Standards state that auditors should consider “subjective as well as objective factors, including management judgment, which is often based on knowledge and experience” (PCAOB AS 2501.03). However, auditing standards do not provide any specific guidance on how to identify and incorporate relevant source information into evidence evaluation. Given the lack of guidance, auditors are unlikely to perform a new, separate assessment of management sources during evidence evaluation when they have already done so during the planning phase of the audit. This would be inefficient, or even impractical at times, given that auditors have limited resources and face substantial budget constraints (Low and Tan 2011; Lopez and Peters 2011). With the significant time delay between the initial consideration of management during planning and fieldwork phases, auditors are likely to focus on the overall conclusion formed instead of the details supporting that conclusion (Kida and Smith 1995; Kida, Smith, and Maletta 1998; Reyna, Corbin, Weldon, and Brainerd 2016).
Audit documentation often reinforces the more holistic conclusion of management’s credibility instead of the specific details supporting the conclusion. Auditing standards require the source of the evidence to be included in audit documentation (PCAOB AS 1215.04). As a result, auditors include statements such as, “Evidence provided by Mike Williams, Accounting Manager” (Walters 2016). These statements document the source of evidence and remind auditors they have previously evaluated the source and concluded he is credible, alleviating the need to perform any additional evaluation of the source during evidence evaluation and documentation. Thus, while auditing standards speak to the importance of the source throughout the audit, it is likely the auditor only explicitly considers the source during audit planning.

**Credibility and Competence**

The terms credibility and competence are often used interchangeably. However, credibility is multidimensional, and competence is one dimension of credibility. Qualities associated with the source such as trustworthiness, likability, competence, and attractiveness are all dimensions of credibility that influence overall source credibility assessments (Pornpitakpan 2004; O’Keefe 2016). A source does not need to be considered high on all dimensions in order to be considered credible from an overall perspective (Pornpitakpan 2004; O’Keefe 2016). Credibility is often viewed from a broad, general perspective, considering the relevant dimensions and forming one overall, holistic assessment of an individuals’ credibility (O’Keefe 2016). But, it can be equally important to explicitly consider specific dimensions of credibility, such as competence, separately. A competent individual is defined as one who has the requisite knowledge or skills for a task. That is, a credible individual may be an expert in one area but lack the necessary knowledge and background in other areas (Pornpitakpan 2004; O’Keefe 2016). Thus, it is important to recognize that competence is context-specific (O’Keefe 2016).
Credibility and competence are important factors when determining the relevance and reliability of audit evidence (PCAOB AS 2301, 2501, 2502). Auditors place more weight on evidence provided by credible sources (Hirst 1994, Anderson et al. 1994) and are able to identify differing levels of credibility and discount information provided by less credible sources (Hirst 1994; Bhattacharjee, Moreno, and Riley 2012). However, by relying on overall credibility assessments made earlier during the audit, auditors may fail to identify and incorporate relevant dimensions of credibility, such as context-specific competence, when evaluating evidence. Further, auditing standards provide auditors no guidance for identifying, and separately considering, these relevant dimensions of credibility. From an auditing standpoint, this is extremely important given the weight auditors’ place on evidence obtained from credible sources and the importance of incorporating relevant source information throughout the audit (e.g., into evidence evaluation and assessment of the reasonableness of the financial statements).

**The Sleeper Effect of the Source**

Both the source and message play an important role in persuasion. Persuasion is highest when both a credible source and a strong message are present (Cook and Flay 1978; Eagly and Chaiken 1993), with both factors increasing the persuasiveness of the message. Strong messages can be discounted when relayed by a source lacking credibility. Conversely, weak messages may be boosted if relayed by a credible source (O’Keefe 2016).

The persuasiveness of a message is generally strongest at the moment of communication and decreases over time. However, this is not always the case. Psychology research has identified scenarios where persuasion increases over time. A recently identified phenomenon known as the “sleeper effect of the source” occurs when the persuasiveness of a weak message increases over time as a result of its association with a credible source (Albarracín et al. 2017).
The sleeper effect of the source highlights the persisting influence of a credible source on persuasion (Albarracín et al. 2017). The term “sleeper effect” is used to describe the temporal changes in persuasion as one aspect of the communication (the message or the source) persists while the other diminishes (Kumkale and Albarracín 2004; Albarracín et al. 2017). The sleeper effect of the source considers a situation where source attributes persist. For the sleeper effect of the source to occur, information regarding the source needs to be accessible during subsequent evaluations of the persuasiveness of the message. The act of documenting the source during evidence evaluation, as required by auditing standards (PCAOB AS 1215.04), reminds auditors they have previously evaluated the source and reinforces their previous (overall) conclusion that management is credible (O’Keefe 2016).

When evaluating subjective audit areas, “the auditor often has to rely on evidence that is persuasive rather than convincing” (PCAOB AS 1015.11). That is, the evidence may not be conclusive, leading the auditor to evaluate it based on its persuasiveness. If, as time passes, source attributes persist and auditors focus on the overall strength of the source more so than the strength of the evidence, they may give relatively unpersuasive evidence provided by a credible source more credence than it warrants. In other words, as time passes, auditors may view weak evidence as more persuasive as a result of its association with a credible source and, in turn, they may fail to identify underlying inconsistencies or potential issues with the audit evidence. As such, we predict that auditors are subject to a sleeper effect of the source: auditors provided weak audit evidence by an overall credible management source will consider evidence to be more

---

7 The sleeper effect has traditionally been examined in cases where persuasive messages are provided by non-credible sources. In this study, we focus solely on the sleeper effect of the source. For more information on the traditional sleeper effect, see Kumkale and Albarracín (2004).

8 The term “evidence” is used when discussing audit information provided by management. In psychology research more broadly, the terms “message” or “argument” are used when discussing the sleeper effect and persuasion.
reasonable when the evaluation of evidence is performed as a separate step later in the audit (as is typically the case) than if the auditor were to evaluate the information regarding both management’s credibility and the evidence concurrently.

**H1a:** When a credible member of management provides weak audit evidence, auditors’ evidence reasonableness assessments will be higher when the assessment of management is made separately (i.e., earlier) compared to when assessments of both the source and the evidence are performed concurrently.

### Context-Specific Competence and the Sleeper Effect

Management credibility assessments made during the planning stages of the audit likely fail to devote much attention to context-specific factors, such as first-hand knowledge and experience with specific audit areas, as they are not yet relevant at this early planning phase. However, when incorporating source information into evidence evaluation, context-specific factors are more relevant than overall credibility as they speak directly to the source’s knowledge and experience with the particular audit area and, thus, ought to influence auditors’ evidence evaluations.

Subjective audit evidence often relies on management judgment, which is “normally based on knowledge and experience” (PCAOB AS 2501.03), and a source’s level of competence can vary by area. For example, assume the accounting manager provides evidence relating to inventory and warranty expense but his background is in inventory. He has no specific experience with warranty expense. If auditors concurrently evaluate the source with the evidence for warranty expense, it will likely be apparent that the source lacks context-specific competence for warranty expense. Given the joint impact of the source and the message on persuasion, the lack of context-specific competence should lead the auditor to consider the evidence for warranty 

11
expense as less persuasive than for inventory. That is, more broadly, when the source and evidence are evaluated concurrently, their joint influence should be readily apparent (because source context-specific factors are considered simultaneously with the evidence they provide), and auditors are likely to assess evidence for an area in which the source has less competence as less persuasive (less “reasonable”) than evidence for an area in which the source has more competence.

However, specific facts about a source’s background are likely to become less accessible as time passes and, as such, auditors are increasingly likely to focus on overall credibility as opposed to more relevant context-specific factors (Kida and Smith 1995; Kida et al. 1998; Reyna et al. 2016). When source and evidence evaluation are temporally separated, auditors’ reliance on overall credibility likely not only ignores context-specific competence, but leads them to treat the source as equally credible across different audit areas for which the source’s competence may vary (e.g., as with the inventory and warranty expense areas in the illustration provided above). As such, auditors’ assessments of evidence provided by an overall credible source are likely to be similar (and relatively high) regardless of the source’s underlying competence in the audit area. Further, and consistent with the main effect predicted by H1a, auditors may view weak audit evidence provided by a credible source as more persuasive as time passes (i.e., a “sleeper effect”).

Taken together, the above discussion suggests that source/evidence assessment timing and context-specific competence interact such that auditors’ evidence reasonableness assessments will be: (a) highest when assessments of the source and the evidence are temporally separated

---

9 Recall that the perceived persuasiveness of a message is a function of both the strength of the source and the strength of the message. Because of the joint influence of the two attributes, a message from a source with context-specific competence is likely to be viewed as more persuasive than one from a source lacking this competence.
(regardless of context-specific competence), (b) lower when assessments of the source and the evidence are concurrent and context-specific competence is high, and (c) lowest when assessments of the source and the evidence are concurrent and context-specific competence is low. Given these expectations, we predict an interaction of the form illustrated in Figure 1.

**H1b:** Source competence and source/evidence evaluation timing will interact such that evidence reasonableness assessments will be highest when the source and evidence are evaluated separately (regardless of the level of context-specific competence), lower when evaluated concurrently and the source has more context-specific competence, and lowest when evaluated concurrently and the source has less context-specific competence.

**Propositional Evaluations: An Intervention**

If auditors’ evaluations of audit evidence are subject to a sleeper effect, there are potential negative implications of relying on management evaluations performed during planning. However, auditors have limited resources and it would be inefficient to perform a separate evaluation or to shift the timing of management evaluations from planning to fieldwork (Low and Tan 2011). Given this, we propose a documentation instruction, based on propositional reasoning, that will help auditors identify and incorporate relevant source information during evidence evaluation without requiring a full reassessment of source credibility during fieldwork.

Propositional reasoning assesses the truth-value of a proposition using logic statements to assess validity of an argument (Johnson-Laird, Byrne, and Schaeken 1992; Gawronski and Bodenhausen 2006). The mere process of evaluating a proposition, whether it is validated or invalidated, results in cognitive elaboration (Gawronski and Bodenhausen 2006). In an audit context, this process will increase one’s attention to relevant source competence characteristics.

---

10 For example, if you want to know if your co-workers John and Tim are knowledgeable about inventory, you may assess this using propositional reasoning: “John works as the Inventory Manager. John has experience with inventory.” Thus, you conclude that John is knowledgeable about inventory. Comparatively, “Tim has no prior experience working with inventory. Tim is the Accounting Manager.” You conclude that Tim has no prior knowledge or experience with inventory and must continue to assess Tim through other means or processes.
Therefore, if audit procedures instruct auditors to explicitly consider whether the source has the necessary context-specific competence relating to the audit evidence area and to document such an assessment, it should reduce reliance on the overall credibility assessment made earlier in the audit and help to counteract potentially negative implications of a sleeper effect. Documentation instructions that include statements to help the auditor directly relate the source’s context-specific competence to the audit area will result in cognitive elaboration (Gawronski and Bodenhausen 2006) and should highlight the source’s context-specific competence (or lack thereof). This will aid auditors in properly identifying and incorporating source context-specific information into evidence evaluation.

**H2:** Given a credible source providing relatively weak audit evidence, auditors’ evidence reasonableness assessments will be lower when audit documentation explicitly instructs the auditor to consider management’s competence as it relates to the audit area during evidence evaluation than when such instruction is absent.

**Examining the Persisting Influence of the Source**

The auditor and management have frequent interactions throughout the audit in order for the auditor to collect relevant information from management (Bennett and Hatfield 2013; 2018). Auditors may revisit a singular client source to gather evidence for multiple audit areas. Arguably, auditors should identify the most appropriate client contact for each audit area and obtain evidence from that individual. However, there are a variety of reasons this may not occur, such as budget pressures, client availability, or existing relationships.\(^{11}\) Given that auditors likely utilize a client contact for a number of audit areas, it is important that auditors are able to identify

---

\(^{11}\) For example, during audit fieldwork, auditors are under time pressure to complete work as efficiently as possible (Sweeney and Pierce 2006). If the most appropriate client contact is not available to discuss an audit area, auditors may discuss the issue with a client manager who is available, which would help the auditor stay on task and finish test-work in an efficient manner. Auditors may also establish relationships with specific client contacts and feel more comfortable seeking out those individuals rather than reaching out to a new client contact, whom they may feel uncomfortable with, when needing to discuss an audit area (Bennett and Hatfield 2013). These factors can lead to repeated experiences with the same client contact, regardless of the source’s competence for the audit area.
the source’s context-specific competence with different audit areas and appropriately incorporate that information into evidence evaluation. Yet, it is unknown how the influence of a credible source, without consideration of context-specific competence, will persist and impact subsequent audit areas, especially if a sleeper effect is present.

Failing to consider context-specific competence will likely influence evidence evaluation in different ways depending on the order in which evidence is obtained. When a source first provides evidence regarding an area in which they have the requisite knowledge and experience, it may serve to strengthen or confirm the initial view that the source is credible. Given that the source was credible for the first area, the auditor’s conclusion that the source is credible may persist and lead the auditor to continue to consider the source as credible for subsequent areas, perpetuating the sleeper effect. Alternatively, when auditors have previously identified a source as being appropriate for an audit area, subsequent areas may be subject to a contrast effect, making a lack of context-specific competence more apparent and counteracting the sleeper effect (Bhattacharjee, Maletta, and Moreno 2007). Thus, we pose a research question investigating whether the order that auditors evaluate audit areas, in which management’s context-specific competence differs, influences the occurrence, and magnitude, of a sleeper effect.

**RQ:** Does the sleeper effect influence auditors’ subsequent assessments of evidence in different audit areas?
III. METHOD

Participants and Experimental Overview

Ninety-three full-time staff auditors and interns completed the study. Participants have an average of nine months of experience working as an auditor.\textsuperscript{12,13} Participants assume the role of staff auditor on a fictitious audit engagement and are provided background information for the client. Participants complete three audit tasks (excluding concurrent evaluation conditions—see below): performing an assessment of management and evaluating audit evidence for two different subjective audit areas.\textsuperscript{14}

Independent Variables

We manipulate source assessment timing and context-specific competence to test the sleeper effect. Source assessment timing is manipulated by having participants either: (a) assess management prior to, and separate from, evidence evaluation, as is common in the current audit environment (separate evaluation condition), or (b) assess management at the same time as the audit evidence (concurrent evaluation condition). Context-specific competence is manipulated through the evidence evaluation task (relating to either inventory obsolescence or plant, property, and equipment impairment (PPE)) being performed. The accounting manager either has more (inventory obsolescence) or less (PPE) first-hand knowledge and experience with the area being audited. This creates a 2×2 between-participants design (source assessment timing × context-

\textsuperscript{12} We utilize lower-level auditors as they frequently gather and evaluate information obtained from management and are considered the most hands-on members of the audit team (Bennett and Hatfield 2013). They are often viewed as the “first line of defense” when it comes to identifying potential audit issues (Bamber 1983). A failure to identify inconsistencies in evidence or properly question audit evidence at the preparer level may reduce the ability of higher-level auditors to identify audit issues (Agoglia, Hatfield, and Brazel 2009).

\textsuperscript{13} The audit staff recruited have an average of 10.3 months of experience. Audit interns have all completed an extended audit internship (through busy season) and have an average of five months of audit work experience. All participants work(ed) for a large national or international accounting firm. Conclusions are unchanged when controlling for months of experience or position (intern versus full-time staff).

\textsuperscript{14} Institutional Review Board approval was granted by the university at which the study was conducted.
specific competence). We also include two cells that allow us to examine the effectiveness of our documentation intervention. Finally, as our participants are asked to evaluate evidence relating to two subjective audit areas, we include two additional cells that manipulate the order in which the two audit areas are presented. As a result, we have a $2 \times 2 + 2 + 2$ design. We utilize the $2 \times 2$ to test H1, the $+2$ documentation intervention cells to test H2, and the $+2$ order cells to examine our RQ. Figure 2 presents the flow of the experiment.

**Task and Procedures**

**Separate Evaluation Conditions**

All participants are provided with background information about the audit client and their role on the audit team. In the separate evaluation conditions, participants first perform an assessment of management. Participants receive information in a “discussion” format with the accounting manager, Mike Williams, about his background and are then asked to assess his competence. The information relates to Mike’s educational background, prior work experience, various roles held within the company, tasks and responsibilities in his current role, and technical qualifications (e.g., licensed CPA). Participants perform the assessment by organizing the information learned about the accounting manager into four broad, but distinct, categories: (1) Areas of Accounting Expertise and Experience, (2) General Accounting Knowledge or Qualifications, (3) Industry Specific Factors, and (4) Other Relevant Information. Participants do not have access to the list of source information after completing the initial assessment.

---

15 The accounting manager’s knowledge and experience are held constant in all conditions. Information provided regarding the accounting manager includes information that speaks to his overall credibility as a source, as well as information regarding his specific competencies.

16 To ensure that participants attend to the information and are not mindlessly categorizing it, they must correctly include three of the four pieces of information regarding his knowledge and experience with inventory in Category (1) before being able to proceed. If participants do not correctly identify and categorize information regarding specific knowledge and experience with inventory, they receive a “review note” from their senior (i.e., an electronic message), instructing them to “revisit the Areas of Accounting Expertise and Experience” category and to “consider relevant work experience with specific accounting areas.”
After completing the source evaluation, the subsequent audit tasks involve participants obtaining audit evidence (via a narrative) from the same accounting manager. Participants evaluate evidence relating to two subjective audit areas. For each audit area, participants obtain seven pieces of audit evidence, three of which are potentially negative (i.e., unconvincing or lacking persuasiveness), indicating the client’s stated balance may not be reasonable.

Once participants finish gathering audit evidence, they complete a filler task in which they are asked to provide CPA exam tips for students planning to take the CPA exam. The filler task separates the evaluation of evidence from prior tasks, which is necessary in evaluating a sleeper effect of the source on auditors’ evaluations of audit evidence. After completing the filler task, staff auditors evaluate the account balance in two ways: a free response conclusion and a scaled response.

**Concurrent Evaluation Conditions**

In the *concurrent evaluation* conditions, the same background information is provided to participants, as well as the same audit tasks to be performed: assessment of management and evaluation of audit evidence. In these two conditions, participants are presented with information regarding management and the audit area at the same time, and they complete their assessments of both management and the audit area *concurrently* (rather than separately). While it is unusual

---

17 As described later, we consider order effects by presenting participants with two audit areas in which we manipulate the order of presentation of the two audit areas.

18 The evidence sets for both audit areas were pretested in a separate study to ensure both audit areas were comparable and not viewed as overly reasonable or unreasonable prior to any experimental manipulations. Participants in this separate study were presented with the set of evidence items for the two audit areas with no experimental manipulations present. Participants were asked “In your opinion, how reasonable is management’s estimate for inventory obsolescence [PP&E Impairment]?” on a scale from 0 (not reasonable at all) to 10 (very reasonable). There was no difference in participants’ reasonableness assessments for the two audit areas (means = 5.87 for Inventory Obsolescence versus 5.45 for PPE Impairment, t = 0.820, p = 0.415, two-tailed).

19 The filler task also mimics current auditing practices regarding gathering evidence through inquiry, in which auditors gather information from management and later evaluate the evidence and document their conclusions (Messier, Glover, and Prawitt 2017).
for assessments of management and evidence to occur concurrently on an actual audit
engagement (Sheriff 1998; Messier et al. 2017), we include the concurrent conditions because
they enable us to establish whether a sleeper effect exists for the more real-world practice of
assessing management prior to, and separate from, evidence. Because evaluation of evidence for
a second (later) audit area would, necessarily, be non-concurrent, it is less useful as a meaningful
benchmark against which to assess the sleeper effect and, thus, we do not require concurrent
condition participants to evaluate both audit areas. Therefore, participants in the concurrent
conditions only evaluate evidence for one audit area, resulting in two between-participants
concurrent evaluation cells: one in which the evidence evaluation task relates to the more
competent area (inventory) and one in which the evidence evaluation task relates to the less
competent area (PPE). As a result, these concurrent conditions should only be compared to the
first audit area assessed in the other conditions, specifically: (1) the separate evaluation cell in
which the auditor first evaluates the more competent area (inventory) and (2) the separate
evaluation cell in which the less competent area (PPE) is evaluated first.

Order Cells

We include two additional cells to control for the order of the evidence evaluation tasks in
the separate evaluation conditions. We manipulate order by presenting either inventory (more
competent area) first and PPE (less competent area) second or PPE (less competent area) first
and inventory (more competent area) second. This allows us to evaluate whether the sleeper
effect generalizes to settings where auditors use the same management source to gather evidence
for multiple audit areas in which the source’s competence may vary.


**Documentation Intervention Cells**

The *documentation intervention* is introduced via the audit instructions for the free-response conclusion. In all of the primary $2 \times 2$ *separate* and *concurrent evaluation* conditions, participants are told that “Megan, your senior, has told you best practices for documentation include identifying the source and their position. For example, ‘Evidence provided by John Doe, Accounting Manager.’”\(^{20}\) The *documentation intervention* manipulation includes additional instructions in which Megan specifically asks participants to document the appropriateness of the source with regard to the audit area for which they are providing evidence. They are also given an example: “Evidence provided by John Doe, Accounting Manager, who is knowledgeable about accounts receivable because he was the Accounts Receivable Manager for 2 years and understands the Company’s AR collection history.”

All case materials provided to participants receiving the *documentation intervention* are identical to those in the *separate evaluation* conditions except for the documentation instructions. We choose to examine the *documentation intervention* (i.e., test its effectiveness) in the *separate evaluation* setting since we expect auditors are susceptible to a sleeper effect when they assess management prior to, and separate from, the evaluation of evidence. Specifically, we test our *documentation intervention* using the separate evaluation condition in which the order of presentation of evidence evaluation tasks is the more context-specific competence area (inventory) first, followed by the less context-specific competence area (PPE), resulting in two *documentation intervention* (within-participant) cells.

\(^{20}\) Audit standards require auditors to include the source of the evidence in audit documentation. Current practices utilize statements such as “Audit Evidence provided by Mike Williams, Accounting Manager” (Walters 2016).
Dependent Variables and Process Measures

The primary variable of interest is participants’ reasonableness assessments of the accounting estimate (based on the evidence gathered) for each audit area (inventory obsolescence and PPE impairment), recorded on an 11-point scale from 0 (Not Reasonable at All) to 10 (Very Reasonable). Participants are also asked how comfortable they are with management’s explanation on a scale from 0 (Very Uncomfortable) to 10 (Very Comfortable); how much advice they would seek out from their senior from 0 (None) to 10 (A Lot); and how likely they are to suggest that the amount may need to be adjusted from 0 (Not at All) to 10 (Very Likely). We also examine participants’ free-response conclusions regarding each audit area.21

After completing audit evidence assessments for both audit areas, participants answer post-experimental questions. Given that the sleeper effect of the source relies on the source being perceived as credible, participants answer questions regarding overall credibility as well as context-specific competence. Participants also answer questions regarding common factors influencing credibility, including likability and trustworthiness of the source, affect toward the source, and ambiguity of the evidence. Finally, participants respond to a series of demographic questions and questions relating to their work experience.

IV. RESULTS

Manipulation Check

As discussed, competence is context-specific and “first-hand experience and knowledge” are critical competence factors in evaluating management and subjective audit evidence (PCAOB AS 2501.03; PCAOB Staff Audit Practice Alert No. 10). Participants are asked to rate

21 One researcher, blind to experimental condition, and one research assistant, blind to hypotheses and experimental condition, independently coded responses for (a) the number of distinct evidence items and (b) the number of relevant source facts. The coders’ initial agreement was 96.1 percent. Cohen’s kappa, a measure of interrater agreement, was 0.862 (values above 0.80 indicate strong to near perfect agreement). Coders resolved all differences.
“Mike’s first-hand experience and expertise” with each of the two audit areas on a scale from 0 (None) to 10 (A Lot). Consistent with the context-specific competence manipulation, participants rated Mike’s first-hand experience and expertise with inventory as significantly higher than with PPE (means = 8.29 and 3.32, respectively; \( t_{1,147} = 17.111, \ p < 0.001 \)), and the results do not differ significantly for separate or concurrent evaluation of management and evidence.\(^{22,23}\)

**Hypothesis Test Results**

**H1: The Sleeper Effect of the Source**

We predict a sleeper effect of the source such that auditors will assess evidence as more reasonable when they assess management competence separately (as opposed to concurrently) from evidence evaluation (H1a), and that this difference is greater when a source is less competent for the specific audit area than when a source has more context-specific competence (H1b). Consistent with the main effect predicted by H1a, we find that participants who evaluate management and evidence separately have higher evidence reasonableness assessments than those who evaluate them concurrently (\( F_{1,71} = 12.470, \ p = 0.001; \) Table 1). With respect to H1b, Figure 3 suggests a good visual fit with our predicted pattern. We test the specific pattern using the following contrast weights: -1 for the more competent/concurrent cell, +2 for more competent/separate cell, -3 for less competent/concurrent cell, and +2 for less competent/separate cell. Table 1, Panel C shows results of this test are significant (\( F_{1,71} = 16.036, \ p < 0.001; \) see Panel D for related pairwise comparisons), and the between-cells residual variance not explained by the contrast is small and not significant (\( q^2 = 0.0288; \) semi-omnibus \( F = 0.321, \ p = 0.727)\),

\(^{22}\) Reported \( p \)-values are one-tailed to reflect directional predictions, unless otherwise noted.

\(^{23}\) Within each condition (where participants see both audit areas), first-hand experience with Inventory is rated as significantly higher than PPE at \( p < 0.001 \). Additionally, across conditions, participants’ ratings of first-hand experience for each audit area do not vary. That is, participants’ ratings of first-hand experience with inventory and PPE are consistent across all conditions (Inventory \( p = 0.900 \) and PPE \( p = 0.706 \), two-tailed).
supporting H1b (Guggenmos, Piercey, and Agoglia 2018). Taken together, these results suggest that auditors are subject to a sleeper effect, and it is more pronounced when client personnel do not have context-specific competence with the audit area.

Additionally, when management and evidence are evaluated separately (compared to concurrently), auditors are more comfortable with management’s explanation, less likely to ask for advice from their senior, and less likely to suggest the account may need to be adjusted (see Table 2). These findings are consistent with H1a. Further, we find that these differences are greater when the source has less context-specific competence, consistent with H1b (see Table 2). If auditors are less likely to discuss evidence provided by management with more senior members of the audit team, audit quality could be compromised.

**H2: Documentation Intervention**

Hypothesis 2 predicts a moderation of the sleeper effect through the use of documentation instructions to explicitly consider a source’s appropriateness for the audit area. The intervention is designed to draw attention to the source’s relevant experience with and knowledge of the audit area, reducing the influence of the sleeper effect of the source. We test H2 with a repeated measures ANOVA, finding a significant main effect. Specifically, the documentation intervention lowers evidence reasonableness assessments for both more and less competent areas when evaluated separately from management ($F_{1,35} = 7.467, p = 0.010$; Table 3).

We further examine the documentation intervention by analyzing the more-competent audit area individually and including comparisons to the concurrent evaluation condition to evaluate the effectiveness of the intervention. Using planned contrasts, we find that when auditors receive

---

24 This test uses the within-participants conditions for separate evaluation and separate evaluation—documentation intervention in which context-specific competence is manipulated within participants (participants evaluate both audit areas with the more competent area first and the less competent area second) and the intervention is manipulated between participants.
the documentation intervention, they assess evidence reasonableness significantly lower than when not receiving the intervention (means = 6.17 and 7.47, respectively, p = 0.005; Table 4 and Figure 4) and there is no difference from those in the concurrent evaluation condition (means = 6.17 and 6.12, respectively, p = 0.940, two-tailed; Table 4 and Figure 4). In addition to lower reasonableness assessments, participants provided with the documentation intervention are less comfortable with management’s explanation, indicate a higher likelihood of seeking advice from their senior, and indicate a higher likelihood of suggesting the account may need to be adjusted (see Table 5). We also find that these measures are no different from the concurrent evaluation condition, providing further support for the effectiveness of the documentation intervention.

Finally, we examine the total number of evidence items and relevant source facts documented in participants’ free response conclusions of the audit evidence. The documentation intervention increases the total pieces of evidence documented during evidence evaluation (intervention present = 2.06 and intervention absent = 0.42, Z = 3.64, p < 0.001; see Table 4 for means). Further, this increase in items is driven by the number of relevant source facts documented in participants’ free response conclusions of the audit evidence. We also examine the documentation intervention for the less competent area; however, we are limited in our analysis due to the experimental design. The documentation intervention is manipulated for one within-participant cell, where the less competent area is evaluated after the more competent area. When the less competent area is evaluated after the more competent area (intervention absent), the initial sleeper effect is dampened due to a contrast effect (evidence reasonableness means: concurrent evaluation = 4.75 and separate evaluation = 5.26, p = 0.272; see the Research Question discussion below for more details on order effects). However, the results for the less competent area are consistent with the more competent area. The presence of the intervention lowers mean evidence reasonableness assessments (4.00 versus 5.26 without the intervention, p = 0.028) and these lower assessments are more in line with the concurrent evaluation condition (4.00 versus 4.75 for the concurrent evaluation condition, p = 0.359, two-tailed). Participants’ responses are coded by one researcher blind to experimental condition and one research assistant blind to hypotheses and experimental condition, both with auditing experience. Responses are coded for each distinct: (a) reference to evidence items provided by management and (b) mention of context-specific source information. With respect to the latter, a statement such as “evidence provided by the accounting manager, who has extensive prior experience in inventory management because of his prior role as inventory manager” would be coded as relevant source information, while “evidence provided by the accounting manager” would not. We use a Poisson regression to test the contrast as they are used to analyze count data that is often skewed toward low values such as 0 or 1. A non-significant $\chi^2$ indicates there is no evidence of a lack of fit of the model. The models appear to be a good fit (total evidence: $\chi^2 = 37.385$, p = 0.360; source items: $\chi^2 = 12.333$, p = 0.999 ).

---

25 We also examine the documentation intervention for the less competent area; however, we are limited in our analysis due to the experimental design. The documentation intervention is manipulated for one within-participant cell, where the less competent area is evaluated after the more competent area. When the less competent area is evaluated after the more competent area (intervention absent), the initial sleeper effect is dampened due to a contrast effect (evidence reasonableness means: concurrent evaluation = 4.75 and separate evaluation = 5.26, p = 0.272; see the Research Question discussion below for more details on order effects). However, the results for the less competent area are consistent with the more competent area. The presence of the intervention lowers mean evidence reasonableness assessments (4.00 versus 5.26 without the intervention, p = 0.028) and these lower assessments are more in line with the concurrent evaluation condition (4.00 versus 4.75 for the concurrent evaluation condition, p = 0.359, two-tailed).

26 Participants’ responses are coded by one researcher blind to experimental condition and one research assistant blind to hypotheses and experimental condition, both with auditing experience. Responses are coded for each distinct: (a) reference to evidence items provided by management and (b) mention of context-specific source information. With respect to the latter, a statement such as “evidence provided by the accounting manager, who has extensive prior experience in inventory management because of his prior role as inventory manager” would be coded as relevant source information, while “evidence provided by the accounting manager” would not.

27 We use a Poisson regression to test the contrast as they are used to analyze count data that is often skewed toward low values such as 0 or 1. A non-significant $\chi^2$ indicates there is no evidence of a lack of fit of the model. The models appear to be a good fit (total evidence: $\chi^2 = 37.385$, p = 0.360; source items: $\chi^2 = 12.333$, p = 0.999 ).
facts (means = 1.50 and 0.00, Z = 115.69, p < 0.001; Table 4).28 Taken together, the results indicate that the documentation intervention assists auditors in identifying and incorporating relevant source information into evidence evaluation without requiring auditors to sacrifice efficiency by changing the timing of management evaluations or re-performing work completed earlier in the audit.

**Research Question**

Given that auditors frequently revisit a single client contact for multiple audit areas, a single source may provide evidence for areas in which they are both more and less competent. We investigate how previous interactions with the same client contact whose context-specific competencies differ across audit areas can alter auditors’ ability to identify and incorporate relevant source information into evidence evaluation. Specifically, we examine whether the order in which auditors obtain evidence from management with differing levels of context-specific competence influences evidence reasonableness assessments. Using a 2×2 repeated measures ANOVA (Context-Specific Competence (within-participants) × Order), we find a significant interaction of order and competence (F_{1,36} = 21.319, p < 0.001; Table 6). In addition, there is an effect of order of evaluation between participants (F_{1,36} = 4.660, p = 0.038; see Table 6 and Figure 5).

We further examine the effect of order by analyzing the pattern of means for each audit area from first to second evaluation. Reasonableness assessments for more competent areas are no different depending on the order of evaluation (more competent means: first = 7.47 and second = 7.79, t_{1,36} = 0.689, p = 0.495, two-tailed, untabulated), with a sleeper effect occurring

---

28 Results for the effectiveness of the documentation intervention for PPE impairment are similar to those for inventory, with participants documenting more relevant source items (Z = 3.02, p = 0.003, untabulated; model fit $\chi^2 = 43.567$, p = 0.867). Although the number of total evidence items documented for PPE is not significant (Z = 1.59, p = 0.110, untabulated; model fit $\chi^2 = 88.441$, p = 0.280), it is directionally consistent with the inventory result.
for more competent areas despite order of analysis. However, when obtaining evidence from a source who is considered less competent for an audit area when they have previously obtained evidence from the same source for an area in which they are considered more competent, auditors are able to recognize a difference in source competence and incorporate this information into evidence assessments. Evidence reasonableness assessments are significantly lower for the less competent area when it follows the more competent area (less competent area means: first = 7.16 and second = 5.26, \( t_{1,36} = 2.670, p = 0.011, \) two-tailed, untabulated). This suggests that, in some cases, the sleeper effect may be dampened as auditors partially “self-correct”. The ability to self-correct is encouraging. However, auditors are placed under constant time pressure and often complete tasks as they come up (Low and Tan 2011; Lopez and Peters 2011). As a result, they are likely unable to strategically consider how and when to evaluate individual audit areas in order to mitigate a potential sleeper effect. Given this, our results suggest firms should consider the use of a documentation intervention to facilitate auditors’ focus on relevant source competence qualifications throughout the audit.

Additional Analyses

*Context-Specific Competence*

As discussed previously, evaluations of the source are often performed at temporally different points in the audit. Given these temporal differences, certain source attributes may become dissociated from the message while the overall assessment remains. While overall credibility is important, auditing standards suggest that certain source attributes, such as context-specific competence, are more relevant than others during evidence evaluation. Therefore, it is important for auditors to be able to identify and incorporate relevant context-specific
competencies, not just overall credibility and competence, into evidence evaluation. Our documentation intervention is aimed at assisting auditors with this process.

We examine participants’ evaluations of context-specific competence for both audit areas, as well as perceptions of the source’s overall competence and trustworthiness, as these are two of the most common dimensions of overall credibility. Overall perceptions of both competence and trustworthiness do not differ between the intervention present and intervention absent conditions (competence means: present = 7.28 and absent = 7.37 on a 0 (Not Competent at All) to 10 (Very Competent) scale, \( t_{1,35} = 0.207, p = 0.837 \), two-tailed; trustworthiness means: present = 1.83 and absent = 2.42 on a -5 (Very Untrustworthy) to 5 (Very Trustworthy) scale, \( t_{1,35} = 1.159, p = 0.254 \), two-tailed). However, in contrast to assessments of overall credibility, any explicit consideration of management’s context-specific competence should lower these perceptions. Consistent with this notion, context-specific competence is lower when the documentation intervention is present than when it is absent (more competent means: present = 6.72 and absent = 7.68, \( t_{1,35} = 1.963, p = 0.029 \); less competent means: present = 3.06 and absent 4.26, \( t_{1,35} = 1.644, p = 0.050 \)), despite no differences in source information between conditions. Thus, the documentation intervention helps auditors identify and consider the source’s appropriateness for the audit area, separately from overall credibility.

**Sufficient and Appropriate Evidence**

Auditing standards charge auditors with obtaining evidence that is both sufficient and appropriate. Sufficiency is defined as “the quantity of audit evidence” and appropriateness is defined as “the quality of audit evidence” (PCAOB AS 1106.04-06). As such, auditors’ evidence reasonableness assessments should be driven by their perceptions of both its sufficiency and appropriateness. Audit standard 1106 – *Audit Evidence* – specifically refers to the importance of
the source, noting that “evidence depends on the nature and the source of the evidence” (PCAOB AS 1106.08). If auditors are incorporating source information into evaluations of audit evidence, as prescribed by standards, source information should influence perceptions of evidence sufficiency and appropriateness.

Our documentation intervention is designed to aid auditors in incorporating relevant context-specific competence into evidence evaluation as opposed to the overall view of management (“management is credible”) formed during planning. Any explicit consideration of management’s context-specific competence should lead to lower perceptions of context-specific competence compared to overall credibility, as our results suggest auditors are not currently considering context-specific competence. Evidence provided by a manager with more context-specific competence should be viewed as relatively more sufficient and appropriate than evidence provided by a manager with less context-specific competence. We explore whether context-specific competence and the sufficiency and appropriateness of audit evidence mediate the influence of our intervention on evidence reasonableness assessments.

We ask participants about the sufficiency and appropriateness of the audit evidence for each audit area evaluated. Given that auditors are charged with obtaining evidence that is both sufficient and appropriate, we create a composite evidence score for each audit area. We test a structural path model to examine the influence of the documentation intervention on evidence reasonableness assessments through both audit areas (see Figure 6). The model can be viewed

29 Participants were asked “To what extent do you believe the evidence you gathered to support the Inventory Obsolescence Reserve [PPE Impairment] amount is sufficient?” from 0 (Not Sufficient at All) to 10 (Very Sufficient) and “To what extent do you believe the evidence you gathered to support the Inventory Obsolescence Reserve [PPE Impairment] amount is appropriate?” from 0 (Not Appropriate at All) to 10 (Very Appropriate).
30 For both areas, the two questions are highly correlated (more competent area (inventory) r = 0.878, less competent area (PPE) r = 0.960), and factor analysis extracts only one factor for each area.
31 We utilize IBM AMOS for our mediation analysis. AMOS is a statistical program for testing structural path models that allows the researcher to build a custom, or specific, model to be tested.
in two parts: (1) the first order indirect effect of the documentation intervention on evidence reasonableness assessments (i.e., the effect on the first audit area evaluated) and (2) the full indirect effect of the intervention on multiple audit areas evaluated (i.e., the effect on both the first and the second audit areas, within participants). The overall model fits the data well ($\chi^2 = 10.470, p = 0.234$) and accounts for a good proportion of the variance (GFI = 0.930; Kline 2011).

In the first part of the model, we find that the effect of the documentation intervention is mediated by perceptions of context-specific competence and evidence sufficiency and appropriateness. Specifically, for the first audit area evaluated, when the documentation intervention is used, evidence reasonableness assessments are statistically lower, as indicated by a 95 percent confidence interval that does not include zero (95 percent CI [-0.528, -0.101], $p = 0.002$). The documentation intervention first influences auditors’ assessments of context-specific competence ($p = 0.047$) and then their assessments of the sufficiency and appropriateness of the audit evidence ($p = 0.005$), which in turn jointly influence auditors’ evidence reasonableness assessments for the first audit area (see Figure 6).

The second part of the model examines the persisting influence of the documentation intervention when multiple audit areas are evaluated. We find that the documentation intervention (relative to no intervention) not only lowers evidence reasonableness for the first area evaluated, but also influences subsequent evaluations, lowering evidence reasonableness assessments for the second area evaluated. The auditors’ evaluation of the second audit area is influenced by both the documentation intervention and the previous audit area evaluated. The full indirect path model is significant (95 percent CI [-0.418, -0.037], $p = 0.009$). The full model highlights the effectiveness of the documentation intervention, even when an earlier assessment (in a different audit area) influences subsequent audit areas.
V. CONCLUSION

We explore whether credible management sources bias auditors’ evidence evaluations. We find that auditors are susceptible to a sleeper effect in which they over-rely on weak audit evidence when it is provided by a generally credible source. Further, we find that current auditing practices can lead auditors to ignore management’s context-specific competence during evidence evaluation, exacerbating the sleeper effect. We also identify a documentation-based intervention that assists auditors in identifying and incorporating context-specific competence into evidence evaluation. Our intervention requires auditors to document a source’s appropriateness for the audit area they are evaluating, which results in (a) lower evidence reasonableness assessments, (b) lower levels of comfort with management’s explanation, and (c) a higher likelihood of bringing potential issues to the attention of a more senior member of the audit team. Collectively, our documentation intervention makes it more likely that weak audit evidence will be examined more thoroughly. We also test a mediation model and find that our documentation intervention first assists auditors in identifying relevant source attributes and then, more importantly, aids integration of this information into auditors’ assessments of evidence sufficiency and appropriateness.

As with all studies, ours is subject to limitations. For example, we explore a setting in which a generally credible source provides relatively weak evidence for a subjective audit assessment. To increase the generalizability of our findings, future research should explore the influence of the sleeper effect in settings in which task subjectivity, overall source credibility, and evidence strength vary. Additionally, we explore only one dimension of overall credibility: context-specific competence. Future research could explore other dimensions of credibility and
how they influence the persuasiveness of a source’s message over time, which in turn can influence judgments that incorporate information conveyed by that message.

This study makes several important contributions. We contribute to both the auditing and psychology literatures by providing new evidence on the sleeper effect. We find evidence of the occurrence of a sleeper effect in professional domains and demonstrate that the sleeper effect not only influences the persuasiveness of an argument over time, but can also influence one’s judgment and decision making when considering the argument. We also address a gap in both the auditing and the broader psychology literatures on source credibility by highlighting the context-specific nature of competence. Specifically, unlike prior research, our study examines competence as a context-specific construct and explores the influence it has on judgments involving the source of information. We also provide empirical evidence to support the theoretical discussion in the psychology literature regarding context-specific competence.

Finally, our study also provides significant practical insights for audit professionals, standard setters, and regulators. Auditing standards highlight the importance of the source of evidence but provide little guidance on how to evaluate and incorporate relevant source information into evidence evaluation. This study demonstrates how current auditing practices may subject auditors to a sleeper effect. The sleeper effect can lead auditors to over-rely on evidence provided by generally credible members of management, even when management lacks the necessary area-specific competence. Thus, auditors need to be aware of, and audit firms should take steps to avoid (e.g., through guidance and instruction), this tendency to over-rely on generally credible sources. Another important finding for both practitioners and standard setters is that, when auditors do consider context-specific competence, it impacts their evaluations of the sufficiency and appropriateness of audit evidence. While auditing standards indicate the critical
nature of considering context-specific competence when evaluating evidence, it is typically rarely done in practice. As such, standard setters may want to consider incorporating the idea of context-specific competence into the assessment of audit evidence. We also provide audit firms with a simple documentation instruction that can be easily (and immediately) implemented to help auditors identify and incorporate context-specific source information into evidence evaluation, mitigating potential negative effects of a sleeper effect.
REFERENCES


Public Company Accounting Oversight Board (PCAOB) Due Professional are in the Performance of Work, Auditing Standard 1015.
Public Company Accounting Oversight Board (PCAOB) *Audit Evidence*, Auditing Standard 1105.

Public Company Accounting Oversight Board (PCAOB) *Audit Documentation*, Auditing Standard 1215.

Public Company Accounting Oversight Board (PCAOB) *Audit Planning*, Auditing Standard 2101.


Public Company Accounting Oversight Board (PCAOB) *Auditing Accounting Estimates*, Auditing Standard 2501.

Public Company Accounting Oversight Board (PCAOB) *Auditing Fair Value Measurements and Disclosures*, Auditing Standard 2052.


**Figure 1**

**H1b Prediction: Interaction of Source Assessment Timing and Context-Specific Competence**

<table>
<thead>
<tr>
<th>Evidence Reasonableness Assessments</th>
<th>Concurrent Evaluation</th>
<th>Separate Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>More Context-Specific Competence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less Context-Specific Competence</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Variable Definitions:**

*Evidence Reasonableness Assessments:* Auditor assessments of reasonableness of the account balance being tested based on the evidence obtained.

*Context-Specific Competence:* Relates to the level of competence the management source has for the audit area being evaluated. The management source either has direct experience and knowledge (*more competence*) or little direct experience and knowledge (*less competence*) in the area being evaluated.

*Source Assessment Timing:* Relates to the timing of the management source and audit evidence assessments. The two assessments are either performed at the same time (*concurrent evaluation*) or are temporally separated (*separate evaluation*).
**Concurrent Evaluation Conditions:** Two separate conditions, one for each audit area. Participants only evaluate one audit area.

**Separate Evaluation Conditions:** Both audit areas are included, and participants evaluate each audit area. Depending on condition, participants will either evaluate the more competent area first, followed by the less competent area, or vice versa.
H1 Results: The Sleeper Effect of the Source

Variable Definitions:

Evidence Reasonableness Assessments: Participants’ assessments of the reasonableness (based on the evidence) of the account balance being tested on an 11-point scale from 0 (Not Reasonable At All) to 10 (Very Reasonable).

Context-Specific Competence: Manipulated through the audit task and audit evidence being evaluated, with the source either having more or less first-hand knowledge and experience with the area. All source information is held constant across all conditions.

- More (Less) Context-Specific Competence: The audit task relates to Inventory (PPE), an area in which the client source (i.e., accounting manager) has direct experience and knowledge (little direct experience and knowledge), leading to more (less) competency in that area.

Source Assessment Timing

- Concurrent Evaluation: There is no temporal delay between assessments; auditors evaluate management and the audit evidence at the same time.
- Separate Evaluation: There is a temporal delay between assessments; auditors evaluate management before, and separately from, their audit evidence evaluation.
Variable Definitions:

Evidence Reasonableness Assessments: Participants’ assessments of the reasonableness (based on the evidence) of the account balance being tested on an 11-point scale from 0 (Not Reasonable At All) to 10 (Very Reasonable).

Source Assessment Timing

- Concurrent Evaluation: There is no temporal delay between assessments; auditors evaluate management and the audit evidence at the same time.
- Separate Evaluation: There is a temporal delay between assessments; auditors evaluate management before, and separately from, their audit evidence evaluation.
- Separate Evaluation—Documentation Intervention: There is a temporal delay between assessments; auditors evaluate management before, and separately from, their audit evidence evaluation. The documentation intervention explicitly instructs auditors to document the source’s appropriateness for the audit area.
Variable Definitions:

Evidence Reasonableness Assessments: Participants’ assessments of the reasonableness (based on the evidence) of the account balance being tested on an 11-point scale from 0 (Not Reasonable At All) to 10 (Very Reasonable).

Context-Specific Competence: Manipulated through the audit task and audit evidence being evaluated, with the source either having more or less first-hand knowledge and experience with the area. All source information is held constant across all conditions.

- More (Less) Context-Specific Competence: The audit task relates to Inventory (PPE), an area in which the client source (i.e., accounting manager) has direct experience and knowledge (little direct experience and knowledge), leading to more (less) competency in that area.

First Evaluation: Evidence reasonableness assessment for the first audit area evaluated.

Second Evaluation: Evidence reasonableness assessment for the second audit area evaluated.
Figure 6
Mediating Relationship

Panel A: Path Model

Documentation Intervention: Present versus Absent (DI)

First Area: Context-Specific Competence (1stCSC)

First Area: Evidence Sufficiency & Appropriateness (1stSA)

Second Area: Context-Specific Competence (2ndCSC)

Second Area: Evidence Sufficiency & Appropriateness (2ndSA)

First Area: Reasonableness Assessment (RA 1)

Second Area: Reasonableness Assessment (RA 2)

a = -0.962, p = 0.047

b = 0.393, p = 0.005

c = 0.655, p < 0.001

d = 0.506, p = 0.022

e = 0.776, p < 0.001

f = 0.722, p < 0.001

g = -1.09, p = 0.010

h = -0.055, p = 0.720

i = 0.313, p = 0.039

j = -0.124, p = 0.347

k = -0.399, p = 0.373

l = 0.139, p = 0.347

m = 0.051, p = 0.905
### Panel B: Tests of Indirect Paths

<table>
<thead>
<tr>
<th>Path Tested</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First order effect of Documentation Intervention:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DI $\rightarrow$ 1st CSC $\rightarrow$ 1st SA $\rightarrow$ RA 1  \newline (a * b * c)</td>
<td>-0.528, -0.101</td>
<td>0.002</td>
</tr>
<tr>
<td><strong>Full Indirect Path:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DI $\rightarrow$ 1st CSC $\rightarrow$ 1st SA $\rightarrow$ RA 1 $\rightarrow$ 2nd CSC $\rightarrow$ 2nd SA $\rightarrow$ RA 2 \newline (a * b * c * d * e * f)</td>
<td>-0.418, -0.037</td>
<td>0.009</td>
</tr>
</tbody>
</table>

**DI:** *Documentation Intervention* = Participants receive documentation instructions explicitly requiring them to document the source’s appropriateness for the area being audited.

**CSC:** *Context-Specific Competence scores* = Participants’ perception of the management source’s context-specific competence relating to the specific audit area, recorded on an 11-point scale from 0 (None) to 10 (A Lot).

**SA:** *Evidence Sufficiency and Appropriateness scores* = A composite measure of participants’ ratings of evidence sufficiency and appropriateness on 11-point scales from 0 (Not at All) to 10 (Very).

**RA:** *Reasonableness Assessment scores* = Participants’ assessment of the reasonableness (based on the evidence) of the account balance being tested on an 11-point scale from 0 (Not Reasonable at All) to 10 (Very Reasonable).
Table 1
H1: Evidence Reasonableness Assessments

Panel A: Descriptive Statistics (Between-Participants)

<table>
<thead>
<tr>
<th>Context-Specific Competence</th>
<th>Source Assessment</th>
<th>Target Audience</th>
<th>Timing Evaluation</th>
<th>Concurrent Evaluation</th>
<th>Separate Evaluation</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>More Competence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6.12 (2.21)</td>
<td>7.47 (1.43)</td>
<td>6.83 (1.93)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>n = 17</td>
<td>n = 19</td>
<td>n = 36</td>
</tr>
<tr>
<td>Less Competence</td>
<td></td>
<td></td>
<td></td>
<td>4.75 (3.02)</td>
<td>7.16 (2.22)</td>
<td>5.92 (2.89)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>n = 20</td>
<td>n = 19</td>
<td>n = 39</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td>5.38 (2.73)</td>
<td>7.45 (1.85)</td>
<td>7.45</td>
</tr>
</tbody>
</table>

Panel B: ANOVA Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sum of Squares</th>
<th>df</th>
<th>F Stat</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source Assessment Timing (H1a)</td>
<td>66.175</td>
<td>1</td>
<td>12.470</td>
<td>0.001</td>
</tr>
<tr>
<td>Context-Specific Competence</td>
<td>13.237</td>
<td>1</td>
<td>13.237</td>
<td>0.119</td>
</tr>
<tr>
<td>Timing × Competence</td>
<td>5.168</td>
<td>1</td>
<td>5.168</td>
<td>0.327</td>
</tr>
<tr>
<td>Error</td>
<td>376.778</td>
<td>71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>465.280</td>
<td>74</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel C: Planned Contrast Test

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>Sum of Squares</th>
<th>df</th>
<th>F Stat</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contrast [-1, +2 -3, +2] (H1b)</td>
<td>85.097</td>
<td>1</td>
<td>16.036</td>
<td>&lt;0.0011</td>
</tr>
<tr>
<td>Residual between-cells variance</td>
<td>3.405</td>
<td>2</td>
<td>0.321</td>
<td>0.727</td>
</tr>
<tr>
<td>Total between-cells variance</td>
<td>88.502</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>376.778</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>465.280</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contrast Variance Residual, q²</td>
<td>0.0288</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 1 (Cont.)

Panel D: Pairwise Comparisons

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>t Stat</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>More Competent: Separate &gt; Concurrent</td>
<td>34</td>
<td>2.213</td>
<td>0.017^1</td>
</tr>
<tr>
<td>Less Competent: Separate &gt; Concurrent</td>
<td>37</td>
<td>2.823</td>
<td>0.004^1</td>
</tr>
<tr>
<td>Separate: More Competent ≠ Less Competent</td>
<td>36</td>
<td>0.522</td>
<td>0.605</td>
</tr>
<tr>
<td>Concurrent: More Competent &gt; Less Competent</td>
<td>35</td>
<td>1.547</td>
<td>0.065^1</td>
</tr>
</tbody>
</table>

^1 p-value reported is for one-tailed test. All other p-values reported are two-tailed.

Variable Definitions

Evidence Reasonableness Assessments: Participants’ assessments of the reasonableness (based on the evidence) of the account balance being tested on an 11-point scale from 0 (Not Reasonable At All) to 10 (Very Reasonable)

Context-Specific Competence: Manipulated through the audit task and audit evidence being evaluated, with the source either having more or less first-hand knowledge and experience with the area. All source information is held constant across all conditions.

• More (Less) Context-Specific Competence: The audit task relates to Inventory (PPE), an area in which the client source (i.e., accounting manager) has direct experience and knowledge (little direct experience and knowledge), leading to more (less) competency in that area.

Source Assessment Timing

• Concurrent Evaluation: There is no temporal delay between assessments; auditors evaluate management and the audit evidence at the same time.
• Separate Evaluation: There is a temporal delay between assessments; auditors evaluate management before, and separately from, their audit evidence evaluation.
Table 2

H1: Additional Measures

Panel A: Descriptive Statistics

<table>
<thead>
<tr>
<th>Context-Specific Competence</th>
<th>Comfort with Management’s Explanation</th>
<th>Likelihood of Asking for Advice</th>
<th>Likelihood of Suggesting an Adjustment</th>
</tr>
</thead>
<tbody>
<tr>
<td>More Competence</td>
<td>6.06 (2.13) n = 17</td>
<td>7.53 (1.61) n = 19</td>
<td>6.53 (2.07) n=17</td>
</tr>
<tr>
<td>Less Competence</td>
<td>4.15 (2.90) n = 20</td>
<td>6.74 (2.56) n = 19</td>
<td>7.30 (1.72) n=20</td>
</tr>
</tbody>
</table>

Panel B: ANOVA Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>df</th>
<th>Sum of Squares</th>
<th>F Stat</th>
<th>p-value</th>
<th>Sum of Squares</th>
<th>F Stat</th>
<th>p-value</th>
<th>Sum of Squares</th>
<th>F Stat</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source Assessment Timing (H1a)</td>
<td>1</td>
<td>76.78</td>
<td>13.700</td>
<td>&lt;0.001</td>
<td>38.798</td>
<td>10.072</td>
<td>0.002</td>
<td>65.653</td>
<td>13.028</td>
<td>0.001</td>
</tr>
<tr>
<td>Context-Specific Competence</td>
<td>1</td>
<td>34.009</td>
<td>6.068</td>
<td>0.016</td>
<td>4.496</td>
<td>1.167</td>
<td>0.284</td>
<td>26.749</td>
<td>13.027</td>
<td>0.024</td>
</tr>
<tr>
<td>Timing x Competence</td>
<td>1</td>
<td>5.853</td>
<td>1.044</td>
<td>0.310</td>
<td>1.465</td>
<td>0.380</td>
<td>0.539</td>
<td>4.904</td>
<td>0.973</td>
<td>0.327</td>
</tr>
<tr>
<td>Error</td>
<td>71</td>
<td>397.912</td>
<td></td>
<td></td>
<td>273.488</td>
<td></td>
<td></td>
<td>357.807</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>74</td>
<td>520.347</td>
<td></td>
<td></td>
<td>320.00</td>
<td></td>
<td></td>
<td>459.92</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Panel C: Planned Contrasts Tests

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>df</th>
<th>Sum of Squares</th>
<th>F Stat</th>
<th>p-value</th>
<th>Sum of Squares</th>
<th>F Stat</th>
<th>p-value</th>
<th>Sum of Squares</th>
<th>F Stat</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Contrast (H1b)</strong>&lt;sup&gt;1&lt;/sup&gt;</td>
<td>1</td>
<td>107.918</td>
<td>19.256</td>
<td>&lt;0.001</td>
<td>45.922</td>
<td>11.922</td>
<td>0.001</td>
<td>91.161</td>
<td>18.089</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Residual between-cells</td>
<td>2</td>
<td>14.516</td>
<td>1.295</td>
<td>0.280</td>
<td>0.59</td>
<td>0.076</td>
<td>0.926</td>
<td>10.952</td>
<td>1.087</td>
<td>0.343</td>
</tr>
<tr>
<td>variance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total between-cells</strong></td>
<td>2</td>
<td>122.434</td>
<td>46.512</td>
<td></td>
<td>273.488</td>
<td>320.00</td>
<td></td>
<td>357.807</td>
<td></td>
<td></td>
</tr>
<tr>
<td>variance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Error</strong></td>
<td></td>
<td>397.912</td>
<td></td>
<td></td>
<td>273.488</td>
<td></td>
<td></td>
<td>357.807</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>520.347</td>
<td></td>
<td></td>
<td>320.00</td>
<td></td>
<td></td>
<td>459.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Contrast Variance</strong></td>
<td></td>
<td>10.68%</td>
<td></td>
<td></td>
<td>0.97%</td>
<td></td>
<td></td>
<td>9.53%</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Residual, q&lt;sup&gt;2&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>1</sup> We use the same contrast weights of [-1, +2, -3, +2] for *Comfort with Management’s Explanation* as with *Evidence Reasonableness Assessments* because the variable is measured on the same directional scale. *Likelihood of Asking for Advice* and *Likelihood of Suggesting an Adjustment* are measured in the opposite direction and, thus, we reverse the contrast pattern to reflect the scales used for these two variables [+1, -2, +3, -2].

<sup>2</sup> p-value reported is for one-tailed test. All other p-values reported are two-tailed.

#### Variable Definitions

*Comfort with Management’s Explanation:* Participants rate their comfort with management’s explanation on an 11-point scale from 0 (Very Uncomfortable) to 10 (Very Comfortable).

*Likelihood of Asking for Advice:* Participants’ rate how likely they are to ask their senior for advice on an 11-point scale from 0 (Not At All) to 10 (Very Likely).

*Likelihood of Suggesting an Adjustment:* Participants’ rate how likely they are to suggest to their senior that the account balance may need to be adjusted on an 11-point scale from 0 (Not At All) to 10 (Very Likely).

*Context-Specific Competence:* Manipulated through the audit task and audit evidence being evaluated, with the source either having more or less first-hand knowledge and experience with the area. All source information is held constant across all conditions.

- More (Less) Context-Specific Competence: The audit task relates to Inventory (PPE), an area in which the client source (i.e., accounting manager) has direct experience and knowledge (little direct experience and knowledge), leading to more (less) competency in that area.

#### Source Assessment Timing

- **Concurrent Evaluation:** There is no temporal delay between assessments; auditors evaluate management and the audit evidence at the same time.
- **Separate Evaluation:** There is a temporal delay between assessments; auditors evaluate management before, and separately from, their audit evidence evaluation.
### Table 3
H2: Documentation Intervention—Evidence Reasonableness Assessments

**Panel A: Descriptive Statistics (Within-Participants\(^1\))**

<table>
<thead>
<tr>
<th>Context-Specific Competence</th>
<th>Source Assessment Timing</th>
<th>Separate Evaluation</th>
<th>Separate Evaluation—Intervention</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>More Competence</td>
<td></td>
<td>7.47 (1.43)</td>
<td>6.17 (1.54)</td>
<td>6.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>n = 19</td>
<td>n = 18</td>
<td>n = 37</td>
</tr>
<tr>
<td>Less Competence</td>
<td></td>
<td>5.26 (2.16)</td>
<td>4.00 (1.68)</td>
<td>4.65</td>
</tr>
<tr>
<td></td>
<td></td>
<td>n = 19</td>
<td>n = 18</td>
<td>n = 37</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>6.37 (2.12)</td>
<td>5.08 (1.93)</td>
<td>7</td>
</tr>
</tbody>
</table>

**Panel B: ANOVA Results—Evidence Reasonableness Assessments**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sum of Squares</th>
<th>df</th>
<th>F Stat</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Within Participants</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Context-Specific Competence</td>
<td>88.549</td>
<td>1</td>
<td>47.080</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Context-Specific Competence × Intervention</td>
<td>0.009</td>
<td>1</td>
<td>0.005</td>
<td>0.946</td>
</tr>
<tr>
<td>Error</td>
<td>65.829</td>
<td>35</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Between Participants</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Documentation Intervention (Present versus Absent)</td>
<td>30.530</td>
<td>1</td>
<td>7.467</td>
<td>0.010</td>
</tr>
<tr>
<td>Error</td>
<td>143.092</td>
<td>35</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^1\) For all observations represented in this table, the order of presentation is: (a) More Context-Specific Competence area first and (b) Less-Context Specific Competence second (i.e., the Less-Context Specific Competence observations are within participants).

**Variable Definitions**

*Evidence Reasonableness Assessments:* Participants’ assessments of the reasonableness (based on the evidence) of the account balance being tested on an 11-point scale from 0 (Not Reasonable At All) to 10 (Very Reasonable).

*Context-Specific Competence:* Manipulated through the audit task and audit evidence being evaluated, with the source either having more or less first-hand knowledge and experience with the area. All source information is held constant across all conditions.

- *More (Less) Context-Specific Competence:* The audit task relates to Inventory (PPE), an area in which the client source (i.e., accounting manager) has direct experience and knowledge (little direct experience and knowledge), leading to more (less) competency in that area.
Table 4
H2: Documentation Intervention—More Context-Specific Competence

Panel A: Descriptive Statistics

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Evidence Reasonableness Assessments</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.12</td>
<td>7.47</td>
<td>6.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.21)</td>
<td>(1.43)</td>
<td>(1.54)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>n = 17</td>
<td>n = 19</td>
<td>n = 18</td>
</tr>
<tr>
<td>Number of Distinct Evidence Items Documented</td>
<td></td>
<td>1.18</td>
<td>0.42</td>
<td>2.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.51)</td>
<td>(0.77)</td>
<td>(1.21)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>n = 17</td>
<td>n = 19</td>
<td>n = 18</td>
</tr>
<tr>
<td>Number of Relevant Source Items Documented</td>
<td></td>
<td>0.41</td>
<td>0.00</td>
<td>1.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.28)</td>
<td>(0.00)</td>
<td>(1.04)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>n = 17</td>
<td>n = 19</td>
<td>n = 18</td>
</tr>
</tbody>
</table>

Panel B: ANOVA Results—Evidence Reasonableness Assessments

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sum of Squares</th>
<th>df</th>
<th>F Stat</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source Assessment Timing</td>
<td>21.832</td>
<td>2</td>
<td>3.592</td>
<td>0.035</td>
</tr>
<tr>
<td>Error</td>
<td>155.002</td>
<td>1</td>
<td>3.039</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>176.833</td>
<td>53</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel C: Planned Contrasts Tests

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>t stat</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Separate Evaluation No Intervention &gt; Intervention (H2)</td>
<td>35</td>
<td>2.675</td>
<td>0.0051</td>
</tr>
<tr>
<td>Separate Evaluation No Intervention &gt; Concurrent Evaluation</td>
<td>34</td>
<td>2.213</td>
<td>0.0171</td>
</tr>
<tr>
<td>Separate Evaluation Intervention ≠ Concurrent Evaluation</td>
<td>33</td>
<td>0.076</td>
<td>0.940</td>
</tr>
</tbody>
</table>

1 p-value reported is for one-tailed test. All other p-values reported are two-tailed.

Variable Definitions
Evidence Reasonableness Assessments: Participants’ assessments of the reasonableness (based on the evidence) of the account balance being tested on an 11-point scale from 0 (Not Reasonable At All) to 10 (Very Reasonable).

Number of Distinct Evidence Items Documented: Count of the number of total unique items documented by participants during their free response evaluation of the account balance.

Number of Relevant Source Items Documented: Count of the number of relevant source facts documented by participants during the free response evaluation of the account balance. This count does not include references such as “Evidence provided by ….” and only counts references to items that would indicate the source has the requisite knowledge of and experience with the audit area.
Table 4 (Cont.)

*Source Assessment Timing*

- **Concurrent Evaluation**: There is no temporal delay between assessments; auditors evaluate management and the audit evidence at the same time.
- **Separate Evaluation**: There is a temporal delay between assessments; auditors evaluate management before, and separately from, their audit evidence evaluation.
- **Separate Evaluation–Intervention**: There is a temporal delay between assessments; auditors evaluate management before, and separately from, their audit evidence evaluation. The documentation intervention explicitly instructs auditors to document the source’s appropriateness for the audit area.
Table 5
H2: Documentation Intervention—Additional Measures

Panel A: Descriptive Statistics

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Comfort with Management’s Explanation</td>
<td>6.06 (2.13)</td>
<td>7.53 (1.61)</td>
<td>5.72 (1.96)</td>
</tr>
<tr>
<td>Likelihood of Asking for Advice</td>
<td>6.53 (2.07)</td>
<td>5.37 (1.93)</td>
<td>6.61 (1.85)</td>
</tr>
<tr>
<td>Likelihood of Suggesting an Adjustment</td>
<td>4.94 (2.25)</td>
<td>3.58 (1.92)</td>
<td>5.44 (2.36)</td>
</tr>
</tbody>
</table>

Panel B: Planned Contrasts Tests

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>t stat</th>
<th>p-value</th>
<th>t stat</th>
<th>p-value</th>
<th>t stat</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Separate Evaluation No Intervention &gt; Intervention (H2)(^1)</td>
<td>35</td>
<td>3.045</td>
<td>0.003(^2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Separate Evaluation No Intervention &gt; Concurrent Evaluation(^1)</td>
<td>34</td>
<td>2.306</td>
<td>0.014(^2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Separate Evaluation Intervention ≠ Concurrent Evaluation</td>
<td>33</td>
<td>0.486</td>
<td>0.630</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Separate Evaluation No Intervention &lt; Intervention (H2)(^1)</td>
<td>35</td>
<td></td>
<td></td>
<td>1.926</td>
<td>0.031(^2)</td>
<td>2.644</td>
<td>0.006(^2)</td>
</tr>
<tr>
<td>Separate Evaluation No Intervention &lt; Concurrent Evaluation(^1)</td>
<td>34</td>
<td></td>
<td></td>
<td>1.686</td>
<td>0.050(^2)</td>
<td>1.959</td>
<td>0.029(^2)</td>
</tr>
<tr>
<td>Separate Evaluation Intervention ≠ Concurrent Evaluation</td>
<td>33</td>
<td></td>
<td></td>
<td>0.123</td>
<td>0.903</td>
<td>0.645</td>
<td>0.523</td>
</tr>
</tbody>
</table>
We use the same directional tests for *Comfort with Management’s Explanation* as with *Evidence Reasonableness Assessments* because the variable is measured on the same directional scale. *Likelihood of Asking for Advice* and *Likelihood of Suggesting an Adjustment* are measured in the opposite direction and, thus, we reverse the contrast pattern.

1 p-value reported is for one-tailed test. All other p-values reported are two-tailed.

**Variable Definitions**

*Comfort with Management’s Explanation:* Participants rate their comfort with management’s explanation on an 11-point scale from 0 (Not Comfortable At All) to 10 (Very Comfortable)

*Likelihood of Asking for Advice:* Participants’ rate how likely they are to ask their senior for advice on an 11-point scale from 0 (Not At All) to 10 (Very Likely).

*Likelihood of Suggesting an Adjustment:* Participants’ rate how likely they are to suggest to their senior that the account balance may need to be adjusted on an 11-point scale from 0 (Not At All) to 10 (Very Likely).
### Table 6

**Research Question: Order of Evaluation**  
**Evidence Reasonableness Assessments**

#### Panel A: Descriptive Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>First</th>
<th>Second</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Order of Evaluation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>More Context-Specific Competence</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.47 (1.43)</td>
<td>7.79 (1.40)</td>
</tr>
<tr>
<td></td>
<td>n = 19</td>
<td>n = 19</td>
</tr>
<tr>
<td></td>
<td>Less Context-Specific Competence</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.16 (2.22)</td>
<td>5.26 (2.16)</td>
</tr>
<tr>
<td></td>
<td>n = 19</td>
<td>n = 19</td>
</tr>
</tbody>
</table>

#### Panel B: Repeated Measure ANOVA Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sum of Squares</th>
<th>df</th>
<th>F Stat</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Within Participants</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Context Specific Competence</em></td>
<td>11.842</td>
<td>1</td>
<td>6.580</td>
<td>0.015</td>
</tr>
<tr>
<td><em>Order × Context-Specific Competence</em></td>
<td>38.368</td>
<td>1</td>
<td>21.319</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Error</td>
<td>64.789</td>
<td>36</td>
<td>1.800</td>
<td></td>
</tr>
<tr>
<td><strong>Between Participants</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Order</em></td>
<td>23.211</td>
<td>1</td>
<td>4.660</td>
<td>0.038</td>
</tr>
<tr>
<td>Error</td>
<td>179.316</td>
<td>36</td>
<td>4.981</td>
<td></td>
</tr>
</tbody>
</table>

**Variable Definitions**

*Evidence Reasonableness Assessments:* Participants’ assessments of the reasonableness (based on the evidence) of the account balance being tested on an 11-point scale from 0 (Not Reasonable At All) to 10 (Very Reasonable).

*Context-Specific Competence:* Manipulated through the audit task and audit evidence being evaluated, with the source either having more or less first-hand knowledge and experience with the area. All source information is held constant across all conditions.

- *More (Less) Context-Specific Competence:* The audit task relates to Inventory (PPE), an area in which the client source (i.e., accounting manager) has direct experience and knowledge (little direct experience and knowledge), leading to more (less) competency in that area.

*Order:* The order in which participants perform the audit tasks. Participants either evaluate the more context-specific competence area first, followed by the less context-specific competence area, or vice versa.