



Gutenberg School of Management and Economics
& Research Unit “Interdisciplinary Public Policy”

Discussion Paper Series

*Professional Skepticism and Auditor
Judgment: Does Trait Skepticism Mitigate the
Recency Bias?*

Christopher Koch, Annette Koehler and Kristina Yankova

December 2016

Discussion paper number 1623

Contact Details:

Christopher Koch
Chair of Corporate Governance and Auditing
Gutenberg School of Management and Economics
Johannes Gutenberg University Mainz
55099 Mainz, Germany

christopher.koch@uni-mainz.de

Annette Koehler
Chair of Accounting and Auditing
Mercator School of Management
University of Duisburg-Essen
47057 Duisburg, Germany

annette.koehler@uni-due.de

Kristina Yankova
Mercator School of Management
University of Duisburg-Essen
47057 Duisburg, Germany

kristina.yankova@uni-due.de

**Professional skepticism and auditor judgment:
Does trait skepticism mitigate the recency bias?**

Prof. Dr. Christopher Koch

Chair of Corporate Governance and Auditing
Gutenberg School of Management and Economics
Johannes Gutenberg University Mainz,
55099 Mainz, Germany
christopher.koch@uni-mainz.de

Prof. Dr. Annette Köhler

Chair of Accounting and Auditing
Mercator School of Management,
University of Duisburg-Essen,
47057 Duisburg, Germany
annette.koehler@uni-due.de

Dr. Kristina Yankova

Mercator School of Management,
University of Duisburg-Essen,
47057 Duisburg, Germany
kristina.yankova@uni-due.de

Professional skepticism and auditor judgment:

Does trait skepticism mitigate the recency bias?

Abstract

In two experiments with experienced auditors, we examine *whether* and *how* trait skepticism mitigates the tendency to place a greater emphasis on the most recently presented information, the so-called recency bias. We measure trait skepticism using the Hurtt scale (Hurtt 2010), manipulate information order, and vary presentation mode across two experiments. When information is presented sequentially (experiment 1), we find that auditors who score higher on the subconstructs of trait skepticism related to evidence examination are less likely to overweight contrasting evidence and thus are less prone to the recency bias. When information is presented simultaneously and the task is complex (experiment 2), we observe that auditors with higher trait skepticism exhibit higher cognitive effort which mitigates the recency bias. Our main contribution is the identification and detailed investigation of trait skepticism as a factor mitigating the recency bias, thereby providing evidence for its behavioral manifestation.

Keywords: Professional skepticism; trait skepticism; auditor judgment; recency

1. Introduction

Professional skepticism is key to the success of the auditing profession. Skeptical auditors are expected to show a higher effectiveness in completing audit procedures (Doucet and Doucet 1996). Hurtt (2010) explains this behavior with the tendency of skeptical auditors to question evidence and management's claims, to require more profound and persuasive audit evidence, and to exhibit higher autonomy and deliberation. Standard setters and regulators around the globe share this view and emphasize the importance of professional skepticism to auditing (IAASB 2012; PCAOB 2012). Professional skepticism is also one of the items of the current IAASB work agenda (IAASB 2014; 2015). Professional skepticism is an individual characteristic which "can be both a trait (a relatively stable, enduring aspect of an individual) and also a state (a temporary condition aroused by situational variables)" (Hurtt 2010, 149). According to Hurtt's (2010) professional skepticism framework, both trait and state skepticism lead to a skeptical mindset, which in turn triggers skeptical behavior.

In this study, we investigate *whether* and *how* trait skepticism mitigates cognitive bias. Hurtt (2010) focuses on trait skepticism, develops a psychometric instrument for capturing its multiple dimensions, and validates the instrument for reliability. However, she does not validate it for its power to predict auditor judgment and behavior. Studies using the Hurtt scale for measuring trait skepticism also provide inconclusive evidence on the effect of trait skepticism on auditor judgment (e.g., Carpenter and Reimers 2013; Popova 2013; Quadackers, Groot, and Wright 2014). This may in part be attributed to the difficulty of disentangling trait skepticism from state skepticism in the context of fraud tasks, because in such settings the effects of trait skepticism might go undetected as a result of the dominance of state skepticism (Robinson, Curtis, and Robertson 2013). In our study, we use a non-fraud task for investigating the effects of trait

skepticism. In particular, we use a going concern assessment task, because fraud usually only plays a relatively minor role in evaluating the ability of an entity to continue as a going concern. Further, auditing standards emphasize that the auditor should follow a balanced approach by not only considering facts and circumstances that cast doubt on the entities' ability to continue as a going concern, but also by taking potentially mitigating aspects into account (ISA 570.A2; SAS 126, AU-C 570.9, A2-A3).

Our study provides an investigation on whether the trait skepticism scale of Hurtt (2010) can predict auditor judgment. In particular, we follow the recommendation of Hurtt (2010, 166) that "it may be determined that specific subconstructs [of the professional skepticism scale] are most useful in predicting auditor behavior". Hurtt (2010, 152) distinguishes between three subconstructs: one subconstruct related to the way an auditor examines evidence (a questioning mind, suspension of judgment, search for knowledge); one subconstruct considering the human aspect (interpersonal understanding), and one subconstruct on the ability of the individual to act on the information obtained (self-esteem and autonomy).

In two experiments with a total of 253 experienced auditors from a Big 4 audit firm in Germany, we investigate the impact of trait skepticism on auditor judgment. We consider both a scenario where auditors update their beliefs step-by-step (experiment 1), akin to the auditor collecting and evaluating evidence, and a scenario where auditors have the possibility to integrate all pieces of evidence at the end of the sequence (experiment 2), akin to the auditor reviewing a working paper where the evidence has been collected. The utilization of these two different scenarios allows us to explore in detail the channels through which trait skepticism may influence auditor judgment in the light of different cognitive demands of the task. We manipulate information order between subjects to investigate the tendency to overweight the pieces of

evidence presented last, the so-called recency effect. This cognitive phenomenon has been recognized as being very pervasive and as representing a serious threat to both the efficiency and the effectiveness of the audit (e.g., Ashton and Kennedy 2002; Krull, Reckers, and Wong-on-Wing 1993; Messier and Tubbs 1994).

In both of our experiments, we observe that auditors are prone to the recency bias and provide evidence that trait skepticism naturally serves as a mitigating factor. When evidence is presented sequentially and beliefs are updated step-by-step (experiment 1), we find that the evidence-related subconstructs of trait skepticism mitigate the recency bias. When auditors are given all pieces of evidence at the end of the sequence in a complex task (experiment 2), we find that trait skepticism mitigates the recency bias via higher cognitive effort. Cognitive effort reduces the recency bias in such a setting by enabling auditors to engage in the cognitively demanding simultaneous and holistic processing of all pieces of evidence. We also replicate other well-known information processing effects that are associated with belief revision including the auditor's heightened sensitivity to negative evidence (e.g., McMillian and White 1993).

Our study contributes to the behavioral auditing literature in several important ways. First, we identify and provide initial evidence for trait skepticism as a naturally occurring, auditor-dependent mitigating factor for the recency bias. Second, we investigate in detail the link between trait skepticism and belief revision in auditing, thereby providing valuable insights into the mechanisms and processes underlying the judgment of auditors. In particular, we are able to identify that trait skepticism and its subconstructs mitigate the recency bias by reducing contrast effects when information is processed sequentially and by being associated with higher cognitive effort when information may be processed simultaneously. These results suggest that the Hurtt scale and its subconstructs have behavioral validity and provide empirical support for the multi-

dimensionality of Hurtt's (2010) notion of trait professional skepticism. Third, we provide new evidence that cognitive effort constitutes a debiasing mechanism also for experienced auditors. Finally, we make use of the large number of observations from experienced auditors derived in our controlled setting to confirm the reliability of the Hurtt scale and to document differences in professional skepticism scores across hierarchical ranks.

The remainder of this paper is organized as follows. In Section II we present the theoretical foundations of the present study, the trait skepticism construct and the belief-adjustment model; and in Section III we discuss the role of trait skepticism in the process of belief revision in auditing. Section IV describes the research design. Section V summarizes the key results. Section VI concludes.

2. Theory and prior literature

Professional skepticism and auditor judgment

The professional skepticism framework developed by Hurtt (2010) describes the relationship between individual characteristics and skeptical behavior. It argues that as an individual characteristic, skepticism can be both a trait and a state. Both trait and state skepticism shape the skeptical mindset of the individual, thereby serving as a mediator, and ultimately influence skeptical behavior. In our study we focus on trait skepticism as a “relatively stable, enduring aspect of an individual” (Hurtt 2010, 149).

Trait skepticism is a multi-dimensional construct. It comprises the following six dimensions: 1) questioning mind; 2) suspension of judgment; 3) search for knowledge; 4) interpersonal understanding; 5) autonomy; and 6) self-esteem (Hurtt 2010). Hurtt (2010, p. 152) categorizes these six dimensions into three subconstructs. The first three dimensions constitute the

evidence-related subconstruct that is, as confirmed by our study, particularly relevant to the area of information processing and belief revision. The fourth dimension, interpersonal understanding, captures the human-related subconstruct, while the fifth and sixth dimension, autonomy and self-esteem, are part of the action-related subconstruct relevant for auditor behavior after information is obtained. Based on all six dimensions, Hurtt developed a 30-item psychometric scale which represents the first instrument that is theoretically designed to measure trait professional skepticism. Prior to the development of the Hurtt scale, no coherent skepticism measure existed and a variety of differently nuanced skepticism proxies were employed in auditing research, including trust, distrust, suspicion, and independence (e.g., Shaub and Lawrence 1996).

Prior research using the Hurtt scale for investigating the effects of trait skepticism on auditor judgment is inconclusive. Peytcheva (2014) finds that trait skepticism improves cognitive performance for students but does not find an effect for auditors. Carpenter and Reimers (2013) observe no effect of trait skepticism on auditors' fraud risk assessment. Popova (2013) also fails to detect a relation between trait skepticism and initial fraud risk assessment, but she notices that relatively more skeptical auditors evaluate additional evidence more critically. Quadackers et al. (2014) observe that trait skepticism can explain auditors' assessment of management-provided explanations in a low-risk environment but that it has little predictive ability in a high-risk environment.¹

In summary, prior studies provide only weak evidence of the influence of trait skepticism on auditor judgment. Our study aims to shed light on the behavioral predictiveness of trait

¹ In an unpublished working paper Hurtt, Eining, and Plumlee (2008) find that auditors with higher trait skepticism identify more contradictions and generate more explanations for potential errors in a work paper review, especially when cues induce some skepticism into the setting.

skepticism using four main strategies. First, we focus on trait skepticism by employing a non-fraud setting where state skepticism should play a minor role to disentangle effects from trait and state skepticism. Second, as traits are linked to cognition (McCrae and Costa, 1996), we use a cognitively demanding task that allows us to shed light on the role of trait skepticism on auditors' cognitive performance (susceptibility to cognitive bias) in information processing and belief revision. The need for gaining a profound understanding of the role of skepticism in auditor belief updating has been highlighted by Nelson (2009). Third, we consider the subconstructs of the Hurtt scale and their match with the task (Bonner 1990). Fourth, we investigate auditor judgment more deeply by investigating auditor information processing and belief-adjustment.

Professional skepticism and belief-adjustment

We examine the effects of trait skepticism on auditor information processing based on the framework on belief adjustment developed by Hogarth and Einhorn (1992). The model assumes that people process information in a manner of sequential anchoring and adjustment to ease cognitive strain and to deal with complexity and information load. The decision-maker starts from an initial position, the anchor, which is then adjusted upwards or downwards in light of new information. The adjustment process is based on an individual's subjective evaluation of the relevance of the piece of information and its contrast to the previously held belief. The updated belief serves then as a new anchor, and the process is repeated until all available information is processed and a final position is reached (Lagnado 2011).²

The belief-adjustment model provides a tractable representation of the nature in which information is received and evaluated during the course of an audit (Cushing and Ahlawat 1996;

² For a detailed discussion of the belief-adjustment model, refer to Bamber, Ramsay, and Tubbs (1997).

Krishnamoorthy, Mock, and Washington 1999). It captures both sequential processing, typical when evidence is received at different times from various sources, and simultaneous processing, typical when evidence is available all at one time, for instance while performing a review of working papers and audit work (Ashton and Kennedy 2002; Gibbins 1984; Knechel and Messier 1990; Rasso 2015).

The link between professional skepticism and belief adjustment has been highlighted, but not empirically tested, by Nelson (2009). Relatedly, Griffith, Hammersley, Kadous, and Young (2015) emphasize that professional skepticism is closely connected to how auditors process information. Cushing (2000, 1) also argues that “skepticism affects auditor behavior primarily by influencing the formation and revision of auditor beliefs” and that “auditors who are skeptical will tend to form different beliefs (...), and therefore will tend to process information (e.g., audit evidence) differently in updating their beliefs than will auditors who are not.” In the present study, we provide theoretical arguments and empirical evidence on the validity of this contention, considering both sequential and simultaneous information processing. As we discuss below, these two modes involve distinct cognitive requirements and behavioral implications, enabling us to disentangle the different channels through which skepticism operates in the context of belief revision. The two modes also map onto two different audit tasks, namely the initial gathering and evaluation of audit evidence and the review of evidence that has been gathered.

3. Hypotheses development

Trait skepticism and order effects when information is presented sequentially

When information is presented sequentially and auditors are asked to state their opinion after each piece of information (step-by-step response mode), auditors process information and

form beliefs sequentially. The belief-adjustment model of Hogarth and Einhorn (1992) describes the sequential processing of information as an adaptive process of anchoring and adjustment, bringing along the risk of order effects. Order effects represent deviations from rational behavior because normatively an individual's beliefs and judgments should only be determined by the content of information (i.e., its diagnosticity, relevance, and reliability) and not by the order in which evidence is presented and processed (e.g., Pennington and Hastie 1986). When decision-makers evaluate a series of mixed evidence sequentially by updating their beliefs after each piece of evidence, the belief-adjustment model predicts a recency bias (see Hogarth and Einhorn 1992, for a detailed discussion). A recency bias means that the total belief revision is biased towards the direction of the evidence received most recently.

Auditing research has shown a great interest in the recency phenomenon as it may have severe implications for both the efficiency and the effectiveness of the audit. Recency effects may affect auditors' search for additional evidence and choice of audit procedures to be performed, placing at risk both the cost and timeliness of the audit. Any focus on the order of information, as opposed to its substance, may reduce the accuracy and quality of the audit (Ashton and Ashton 1988; Messier and Tubbs 1994; Trotman 1998). Consistent with the predictions of the belief-adjustment model, prior research finds that recency is a fairly pervasive and robust phenomenon for circumstances where information is presented sequentially (see Trotman and Wright 2000; Kahle, Pinsker, and Pennington 2005, for extensive reviews of prior research). A sequential presentation of information reflects the incremental nature of most audit judgment tasks involving the initial collection of evidence (Gibbins 1984; Ashton and Ashton 1988). However, audit standards generally do not provide rules or guidelines on the order in which information is to be

processed, leading to the risk of biased judgments. In that context, mitigating factors can play a very important role in ensuring bias-free auditor judgments.

Despite the considerable auditing research conducted on the belief-adjustment model, very little is known regarding the influence of personal factors on the process of belief revision (Kahle et al. 2005). As personal factors are temporally stable and cross-situationally consistent, their exploration promises to shed important light on understanding the determinants and processes underlying belief updating. Addressing this issue, and taking into consideration the inherent link between traits and cognition, we investigate the influence of trait skepticism on the mental constructs and processes underlying auditors' belief adjustment in response to new evidence. The question that we ask in our study is whether trait skepticism mitigates the recency bias. Prior research claims, but does not test, the proposition that under certain conditions professional skepticism may contribute to mitigating cognitive biases (e.g., Fahrer 2012; Hurtt, Brown-Libur, Earley, and Krishnamoorthy 2013).

We develop our hypothesis on the mitigating effects of trait skepticism by considering the role of contrast effects as one main mechanism underlying the recency effect. Contrast effects are an integral part of the descriptive belief-adjustment model of Hogarth and Einhorn (1992) and occur when the weight assigned to a new piece of evidence is greater the more it contrasts with the prior belief. For example, when prior beliefs are negative because negative information has been received before, the decision-maker weights the positive information received at the end more heavily, and vice versa. Contrast effects derive from the salience of contrasting stimuli such as contrasting evidence (Fiske 1980; Maurer and Alexander 1991). Evaluators perceive contrasting evidence to be more salient as it is more surprising. The salience effect is moderated when evaluators engage in a more careful processing of information, resulting in a sufficient allocation

of attention also to those pieces of evidence that are not particularly salient (Strack, Erber, and Wicklund 1982).

We expect that auditors with high levels of trait skepticism are less prone to contrast effect because they tend to evaluate evidence more carefully. This reasoning is based on the evidence-related dimensions of trait skepticism. First, suspension of judgment implies an inclination to postpone belief, judgment, and action until there is sufficient persuasive evidence on which to base a judgment or conclusion. We argue that auditors showing this predisposition will base their judgement on a more holistic perspective, giving a more equal weight to both salient and non-salient pieces of evidence. Second, having a questioning mind involves a disposition to search for a more comprehensive evidential basis on which to build one's judgment and to question everything including the own beliefs. Applied to our scenario, taking on a more comprehensive perspective would imply to consider to a larger degree also non-salient pieces of evidence. Third, search for knowledge entails general curiosity and interest. Auditors with this characteristic tend to seek profound information before rendering a judgment with the consequence that they should sufficiently integrate also those pieces of evidence in their judgment that are relevant but not salient. In contrast, the other subconstructs of trait skepticism and their dimensions (interpersonal understanding, autonomy, self-esteem) are not directly related to the way information is processed. Based on this reasoning that takes into account the notion of Hurtt (2010) that an investigation of subconstructs may be beneficial in predicting and explaining auditor judgment, we formulate the following hypothesis:

HYPOTHESIS 1. *In a setting where evidence is presented sequentially and judgments are required after each piece of evidence, auditors with higher degrees of trait skepticism (and,*

in particular, its evidence-related subconstruct) will exhibit less recency effects as compared to those with lesser degrees of trait skepticism.

Trait skepticism and order effects when information is presented simultaneously

When information is *presented* simultaneously and auditors are only asked to state their opinion at the end (end-of-sequence response mode), auditors may engage in either a sequential or simultaneous *processing* of information. As discussed above, a recency bias may occur when information is processed sequentially. A simultaneous processing of information is an alternative cognitive mode in belief updating. It may occur naturally in audit settings when a reviewer examines all the available evidence at one time. This mode involves an overall assessment of all pieces of evidence which are integrated as a whole in a single belief adjustment at the end. Hence, in contrast to sequential processing where beliefs are incrementally updated after each individual piece of information, an end-of-sequence processing involves a single belief adjustment based on the aggregated impact of evidential stimuli with the consequence that contrast effects and thus recency effects do not occur. Therefore, it is important to consider whether auditors engage in a sequential or simultaneous processing of information when predicting the occurrence of the recency bias in an end-of-sequence response mode.

Hogarth and Einhorn (1992) argue that whether information is *processed* sequentially (i.e., step-by-step) or simultaneously (i.e., end-of-sequence) in an end-of-sequence response mode depends on the complexity of the task. Complexity refers to the cognitive demands of a task and is considered to be a function of both the amount of information provided and the familiarity of the decision-maker with the task (Hogarth and Einhorn 1992). For complex tasks and an end-of-sequence response mode, Hogarth and Einhorn (1992) predict that a sequential processing of

evidence is more likely than a simultaneous processing because it promotes cognitive economy by allowing decision-makers to focus on the interim belief as the running total. As discussed above, sequential processing poses the threat of contrast effects and recency. In contrast, simultaneous processing does not pose the risk of recency effects but is cognitively more demanding because a simultaneous aggregation of all pieces of evidence consumes substantial memory capacity and cognitive resources.

Prior research supports the idea that auditors are prone to the recency bias in settings with an end-of-sequence response mode where information is presented simultaneously if the complexity of the task renders a sequential processing of information likely (Tubbs, Messier, and Knechel 1990; Trotman and Wright 1996). Kennedy (1993) conceptualizes recency as an effort-related bias that can be effectively mitigated via effort-enhancing mechanisms that enable a simultaneous processing of information. Prior research generally supports this notion in the context of externally imposed triggers of effort such as accountability (Kennedy 1993) or documentation requirements (Cushing and Ahlawat 1996). However, it is unclear whether there are also innate effort-enhancing mechanisms that can naturally mitigate recency.

Following Kennedy's notion of recency as an effort-related bias, we posit that trait skepticism might reduce this bias through higher cognitive effort enabling a simultaneous processing of information even in complex, cognitively demanding tasks framed in an end-of-sequence response mode. The link between trait skepticism and cognitive effort is most apparent for the evidence-related dimensions of trait skepticism. First, having a desire to investigate and look deeper shows a motivation to spend additional cognitive effort (search for knowledge). Second, constantly questioning everything instead of accepting things at face value requires cognitive effort (questioning mind). Third, withholding judgment and withstanding the desire for

closure is cognitively demanding (suspension of judgment). Confirming the general link between trait skepticism and cognitive effort, Bhaskar, Majors, and Vitalis (2015) demonstrate that higher degrees of trait skepticism are associated with greater amounts of cognitive resources exerted throughout the audit process.

To summarize, we expect that the trait skepticism and in particular its evidence-related subconstruct is associated with higher cognitive effort, which in turn mitigates the recency bias by enabling simultaneous processing of information even for complex tasks. We posit the following hypothesis:

HYPOTHESIS 2. *In a setting with a complex task where evidence is presented simultaneously and judgments are asked for at the end of the evidence presentation, auditors with higher degrees of trait skepticism (and, in particular, its evidence-related subconstruct) will exhibit higher cognitive effort mitigating the recency bias as compared to auditors with lower levels of trait skepticism.*

4. Research design and method

Participants

Auditors from multiple national offices of a Big 4 audit firm in Germany participated in the two experiments of our study.³ To ascertain that participants possess sufficient task knowledge and experience, we only include persons in our sample who (1) have at least three years of general

³ The case materials were in German to enhance participants' involvement with the case study, to ensure understandability, and to reduce completion time. The two main methods for translating psychological instruments are the back-translation and the independent translation technique (see e.g., Gudmundsson 2009 for a discussion of the advantages and disadvantages of each technique). We translated the Hurtt scale into German using an independent translation technique, namely the parallel blind technique (Behling and Law 2000).

audit experience, and (2) have at least some experience with the going concern task.⁴ These requirements should be sufficient since auditors from the level of audit seniors onwards are regularly involved in the gathering, documentation, and evaluation of evidence related to the going concern task (Ashton and Kennedy 2002).

The first experiment with 181 participants was conducted during four national professional training courses under the supervision of one of the authors.⁵ After excluding participants with incomplete responses on the main variables, insufficient experience or extreme initial beliefs, our final sample consists of 161 usable responses.⁶ Table 1, Panel A, shows that the participants have an average of 10.9 years (SD = 7.6 years; range: 3–32 years) of general auditing experience. Our sample includes 69 assistants and seniors, 64 managers and senior managers, 25 partners and directors, and 3 participants at unspecified ranks (Table 1, Panel B).

The second experiment was conducted during three national professional training courses under the supervision of one of the authors or a research assistant. A total of 110 auditors participated. We use the same exclusion criteria as in experiment 1, resulting in 92 usable responses.⁷ Auditors have an average of 13.1 years of experience (SD = 7.3; range: 4–30 years,

⁴ We classify auditors as having at least some experience with going concern tasks when they self-assess their level of experience with these tasks to be at least 2 on a scale from 1 (no experience at all) to 7 (very high level of experience).

⁵ The data from this first experiment was also used in the dissertation of Kristina Yankova. The dissertation is published as a monograph as is customary and required as part of the German academic system.

⁶ In detail, we exclude six participants with incomplete responses on the Hurtt scale and nine participants without experience with going concern tasks. We also exclude five participants with extremely strong positive ($S_0 = 100$) or negative ($S_0 = 0$) initial beliefs, as these beliefs do not allow for upward or downward belief adjustments, leading to floor or ceiling effects (Hite and Stock 1996). Our results are inferentially similar when not applying these exclusion criteria.

⁷ In detail, we exclude five participants with incomplete responses on the Hurtt scale, four participants with missing data on cognitive effort, and four participants without experience with going concern tasks. We also exclude three participants with extremely strong positive ($S_0 = 100$) or negative ($S_0 = 0$) initial beliefs, as these beliefs do not allow for upward or downward belief adjustments, leading to floor or ceiling effects (Hite and Stock 1996). We exclude two participants indicating their previous participation in experiment 1. Our results are inferentially similar when not applying these exclusion criteria.

Table 1, Panel A). Participants include 18 assistants and seniors, 56 managers and senior managers, 17 partners and directors, and 1 participant at an unspecified rank (Table 1, Panel B).

---- Insert Table 1 here ----

Experimental task

We conduct two experiments that share the same experimental task involving a going concern assessment of a financially distressed firm. We base the case information on an actual audit engagement of a Big 4 audit firm and prepare it in close cooperation with the then incumbent engagement partner. We provide background information on the firm's operations and internal controls, and include the financial statements comprising a balance sheet, an income statement, and an aggregated cash flow statement for the financial year under audit as well as the previous two years.⁸

In the next step, we inform participants about newly acquired relevant information. This additional information is mixed, containing both positive and negative cues with respect to auditor's assessment of the firm's survival. By conducting a pretest with 13 experienced auditors, we ensure that the items of additional information are interpreted as intended. As a further validation, auditors classify each cue in terms of its direction and its relevance as part of the post-experimental questionnaire in the main study. We observe that auditors misclassify positive evidence as negative evidence (or vice versa) at a rate of only 1.8 percent and 0.7 percent in experiment 1 and 2, respectively.

⁸ To ensure the validity, consistency, and plausibility of the research materials, we asked three experienced auditors (two partners and one senior manager) and three professors to critically review the case materials and to provide comments on the research instrument. Based on the feedback obtained, we made minor modifications to the experimental materials. Finally, we performed a pilot-test with five academic staff members with practical experience in auditing. This test confirmed the understandability and reasonableness of the case materials.

We manipulate the order in which the additional information is presented by randomly assigning participants to one of two order conditions. In the first condition, we present positive evidence first and the negative items last (+ –). In the second condition, the order of evidence presentation is the opposite (– +). Consistent with prior research, the presentation order of the positive and negative cues is held constant among the processing conditions (e.g., Hogarth and Einhorn 1992).

In the going concern assessment, participants evaluate the likelihood that the company will continue as a going concern for the next twelve months on a scale from 0 percent (“Business activities will not be continued”) to 100 percent (“Business activities will be continued”). We measure auditors’ initial belief regarding the entity’s going concern likelihood prior to presenting the pieces of additional information (S_0)⁹ as well as after presenting a total of N pieces of additional information (S_N). Our main dependent variable is auditors’ belief revision measured as the difference between the final likelihood assessment of the entity’s continued existence (S_N) and the initial assessment (S_0). A recency effect emerges if participants exhibit greater upward belief revisions in the condition where the positive pieces of evidence are presented last. More formally stated, recency effects appear when $S_N - S_0 (- +) > S_N - S_0 (+ -)$.

After the completion of the task, the participants fill out a questionnaire containing the Hurtt scale, an assessment of the direction and relevance of the additional cues, measures of cognitive effort, demographic information, and some general questions about the case. In line with

⁹ We ask auditors to provide self-generated initial assessments to capture their genuine beliefs and to avoid confounding effects such as anchoring on the externally provided prior beliefs (e.g., Asare 1992; Butt and Campbell 1989).

prior literature, we provide the Hurtt scale after the completion of the experimental case (Peytcheva 2014).

Differences in the setting between experiment 1 and experiment 2

We conduct two experiments that differ mainly with regard to the response mode employed. The first experiment uses a step-by-step response mode in which participants are asked to provide an updated evaluation of the company's likelihood to continue to exist after each piece of additional evidence. This response mode ensures that participants process information and update their beliefs step-by-step. The second experiment uses an end-of-sequence response mode in which participants are only asked to provide their final evaluation after having received all pieces of additional information. This setting gives participants the possibility to process information either step-by-step or at the end of the sequence.

The only other differences between the two experiments concern the number and length of the additional pieces of evidence. For experiment 1, the associated hypothesis (H1) is not conditional on the complexity of the additional evidence. Therefore, in line with most prior research, we use an efficient design with four pieces of additional evidence, two positive and two negative. This design is between medium and high complexity based on the extensive background information and the length of our evidence items (average of 47.5 words, range: 27–75 words).¹⁰ In contrast, the hypothesis referring to experiment 2 (H2) only holds for a task of high complexity because recency is only predicted to occur in an end-of-sequence design when the complexity of the task makes some auditors process the information sequentially even if it is presented

¹⁰ Hogarth and Einhorn (1992, fn. 2) mention a study with an average length of 23 words per item as an example for a study that is between simple and difficult. To provide another comparison, Kennedy (1993) uses a design with eight pieces of evidence with an average length of 13.3 words per item. Based on arguments of task familiarity, she classifies the task as simple for auditors but complex for MBA students.

simultaneously. Therefore, we use eight instead of four pieces of mixed evidence with a higher average length of each piece of evidence (72.5 vs. 47.5 words), justifying a categorization of the task as being of high complexity.

5. Results

Professional Skepticism

We measure the participants' degree of trait professional skepticism using the 30-item Hurtt scale (Hurtt 2010). We use our large dataset from the two experiments to further validate the Hurtt scale and to explore differences across hierarchical ranks. The relation between experience and trait skepticism is of interest to regulators wanting to understand the determinants of trait skepticism (IAASB 2015). In comparison to Hurtt (2010), we have a greater number of observations ($n = 253$ vs. 200) of more highly experienced auditors (average years of experience = 11.7 vs. 4.8) and a more controlled setting (supervised sessions vs. internet sessions).

The internal consistency of the full instrument of 30 items is high and similar to that reported by Hurtt (Cronbach's $\alpha = 0.87$ vs. 0.86). As Hurtt, we use oblique rotation and find that almost all items load on the six theorized factors at an acceptable level (> 0.4).¹¹ Hurtt (2010) observes that two items do not have their highest loading on their theorized construct. Our findings confirm the questionable validity of exactly the same two items. Therefore, we exclude these items from further consideration when deriving our measures for the subconstructs of trait skepticism.¹²

¹¹ One item loads on the theorized factor, but to a less acceptable level. This item is: "I often reject statements unless I have proof that they are true" (questioning mind, loading of 0.33).

¹² The first item is "I enjoy trying to determine if what I read or hear is true." We find that this item does not only load on the theorized construct of questioning mind (0.33) but also on search for knowledge (0.36). In Hurtt (2010), this item only loads on search for knowledge. The second item is "I usually notice inconsistencies in explanations." This item is also theorized to load on questioning mind (0.23), but it loads most strongly on self-esteem (0.47) in

In line with Hurtt (2010), the Cronbach's alphas of five of the six subcomponents are in a very appropriate range from 0.75 to 0.88, indicating high reliability. The sixth subcomponent, questioning mind, has an acceptable reliability (Cronbach's alpha of 0.62).

Table 2 shows that the average total skepticism score is 135.3 (SD = 14.3, range: 71–168). The average scores increase slightly as one moves up the hierarchical ranks from assistant/senior (mean (SD) = 133.3 (14.3)) to manager/senior manager (mean (SD) = 135.5 (14.6)) and to partner/director (mean (SD) = 137.8 (13.8)). However, pairwise comparisons indicate that these differences in the total score between the groups are not significant (p-values > 0.2).¹³ We observe no significant differences across hierarchical ranks for all evidence-related dimensions and for the total score of evidence-related trait skepticism (p-values > 0.4). Significant differences across hierarchical ranks appear only for autonomy and self-esteem with partner/director exhibiting the highest levels compared to all other groups (p-values < 0.1).¹⁴

---- Insert Table 2 here ----

Experiment 1: Sequential information presentation

The first experiment uses a step-by-step response mode, requiring auditors to process the information and to update their beliefs sequentially. We confirm the descriptive validity of the belief-adjustment model for our data using the approach of Bamber et al. (1997).¹⁵ Overall, we

our study. In Hurtt, it loads most strongly on autonomy. Our results are inferentially similar when not excluding these two items.

¹³ In all pairwise comparisons, we report p-values that are calculated using the method of Bonferroni to adjust for multiple testing. The results are robust to using other methods.

¹⁴ Significant differences across ranks are only the following ones: Partner/director have higher levels of autonomy compared to assistant/senior (p-value = 0.019). Partner/director also have higher self-esteem than assistant/senior (p-value = 0.008), and manager/senior manager have higher self-esteem than assistant/senior (p-value = 0.006).

¹⁵ First, the belief-adjustment model assumes that prior beliefs do not affect the relevance assessment of an individual piece of evidence, and we find only a weak correlation ($r = 0.132$ for positive evidence, $r = 0.125$ for negative evidence). Second, it assumes that evidence that is rated to be more relevant has a larger impact on belief revision. Accordingly, we find a positive correlation ($r = 0.473$ for positive evidence, $r = 0.395$ for negative evidence)

observe that auditors assess the negative pieces of evidence to be more relevant than the positive ones (mean (SD) = 5.2 (1.06) vs. 4.9 (1.06), t-test: p-value = 0.003, untabulated). More importantly, we find that auditors' trait skepticism is not correlated with their relevance assessments ($\rho = 0.051$, p-value = 0.521). This finding suggests that trait skepticism does not influence auditors' relevance assessments and provides support for the neutral (as opposed to presumptive doubt) nature of Hurtt's skepticism concept and her scale.

Table 3, Panel A, presents the descriptive statistics. The initial belief of the participants that the company will continue as a going concern does not differ significantly across the two experimental conditions (positive evidence last: mean (SD) = 55.0 (21.3) percent vs. negative evidence last: mean (SD) = 52.7 (20.3) percent) (t-test: $p = 0.472$). Once all four pieces of evidence are processed, auditors are more optimistic about the viability of the firm when the positive evidence is presented last (mean (SD) = 63.4 (20.2) percent, +8.4 percentage points) compared to the condition where the negative evidence is presented last (mean (SD) = 48.7 (19.3) percent, -4.0 percentage points). The difference in belief revision between the two order conditions of 12.4 percentage points is significant (t-test: $p < 0.001$, untabulated), confirming the presence of a recency bias. Figure 1 illustrates the resulting fishtail pattern.

H1 predicts that auditors with higher degrees of the evidence-related dimensions of trait skepticism will exhibit a weaker recency effect. For testing our hypothesis, we set up a regression model with the belief revision ($S_4 - S_0$) as a dependent variable (see Table 3, Panel B). In Model I, we include the order of information presentation as an indicator variable (*Order*), the trait skepticism score continuous variable (*PS*), and the interaction term *Order* * *PS*. To facilitate

between relevance assessments and belief revision when controlling for prior beliefs. Moreover, our analysis confirms recency effect and contrast effects (as discussed below).

interpretation, we mean-center all continuous variables. We observe a non-significant interaction effect (Coef. = -0.24, t-stat = -1.33, p-value = 0.185). As we argued that the evidence-related subconstruct of trait skepticism will be most relevant for the task of evaluating evidence, we investigate the effects of trait skepticism in more detail by splitting it up into its evidence-related subconstruct (*PS_Evid*) and its other subconstructs (*PS_Other*). We include these two variables as main effects and in interaction with *Order* in Model II. We find a marginally significant negative sign for *Order * PS_Evid* indicating that the recency bias is lower for auditors with a higher score on evidence-related trait skepticism (Coef. = -0.68, t-stat = -1.72, p-value = 0.087). The other subconstructs do not interact significantly with order (Coef. = 0.07, t-stat = 0.22, p-value = 0.829).

For illustrating the mitigating effects of evidence-related trait skepticism *PS_Evid*, we perform a median split (see Table 3, Panel A, and Figure 1). For the low *PS_Evid* group, the belief revision is 16.5 percentage points more positive when positive (+11.1 percentage points) rather than negative evidence (-5.4 percentage points) is received last. This figure is highly significantly different from 0 (t-test: p-value < 0.001), indicating a strong recency effect. In contrast, the belief revision is only 7.6 percentage points more positive when positive (+5.0 percentage points) rather than negative evidence (-2.6 percentage points) is received last for the high *PS_Evid* group. This figure is still significantly different from 0 (t-test: p-value = 0.041) but smaller than for the low *PS_Evid* group.

For understanding the mechanisms underlying the mitigating effects of evidence-related trait skepticism, we investigate its relationship with contrasts effects. Contrast effects can be measured as the difference in belief revision for the same type of evidence when comparing the scenario where it is presented last to the scenario where it is presented first. For example, for the condition where positive evidence is presented last, contrast effects are the difference between S_4

– S_2 (– +) and $S_2 - S_0$ (+ –). As shown in Table 3, Panel A, and consistent with our finding of recency, auditors in the low *PS_Evid* group have a higher belief revision when positive evidence is presented last (+30.3) than when it is presented first (+21.6), resulting in a contrast effect of 8.7, which is significantly larger than 0 (t-test: p-value = 0.028). In comparison, auditors with high *PS_Evid* only exhibit a contrast effect of 6.8 (26.5-19.7) that is only marginally significantly different from 0 (p-value = 0.082). For negative evidence, the low *PS_Evid* group shows again a significant contrast effect (-7.8, t-test: p-value = 0.015), while the high *PS_Evid* group continues to exhibit no significant contrast effects (-0.8, p-value = 0.832).

Finally, we explore whether other demographic variables, especially auditor experience, have a debiasing effect. In untabulated tests, we find that the interaction effect between *Order* and individual demographic factors is insignificant,¹⁶ except for audit experience in the client industry featured in this case. We elicit industry experience on a scale from 1 (no experience at all) to 7 (very high level of experience) by asking “How would you evaluate your experience in the steel industry?” in the post-experimental questionnaire. Industry experience (*ExpInd*) has a significant interaction effect with *Order* (Coef. = -3.96, t-stat = -2.19, p-value = 0.030, Table 3, Panel B, Model III). This finding indicates a debiasing effect of this type of experience, confirming prior literature (e.g., Trotman and Wright 1996). Including the effects of industry experience into the model does not alter the debiasing effect of evidence-related trait skepticism (Coef. = -0.75, t-stat = -1.87, p-value = 0.063).

--- Insert Figure 1 here ---

--- Insert Table 2 here ---

¹⁶ Our results regarding evidence-related trait skepticism remain inferentially unchanged when including additional demographic variables in the regression model.

In summary, the evidence shows that the evidence-related subconstruct of trait skepticism has a mitigating effect for the recency bias via smaller contrast effects. That is, auditors with high evidence-related trait skepticism overreact less to the contrasting evidence presented last.

Experiment 2: Simultaneous information presentation

The second experiment employs an end-of-sequence response mode. After auditors express their initial belief based on the background information (S_0), they receive all eight pieces of evidence before being asked to state their final belief (S_8). This design enables auditors to process the information and to update their beliefs either in a step-by-step or an end-of-sequence manner. As in the previous experiment, we observe that auditors assess the negative pieces of evidence to be more relevant than the positive ones (mean (SD) = 5.3 (0.82) vs. 4.8 (1.00), t-test: p-value = 0.002, untabulated).

Table 4, Panel A, shows the descriptive statistics. Auditors slightly revise their beliefs upwards when positive evidence is presented last (+0.8 percentage points) but strongly revise it downwards when negative evidence comes last (-10.9 percentage points). The difference of 11.7 percentage points between both order conditions is highly significant (t-test: p-value < 0.01), revealing a recency effect as expected.¹⁷

We use a path model to test our hypothesis on the indirect effect of trait skepticism on recency via cognitive effort (H2). We implement the path model using regression analysis (Table

¹⁷ Kennedy (1993) also uses an end-of-sequence design with eight pieces of evidence but does not observe a recency effect for auditors. To reconcile our findings with her findings, please note that our study uses a complex task with an extensive background story and long pieces of evidence (average = 72.5 words), while her study uses a simple task with no background information and short pieces of evidence (average = 13.3 words). Compared to Kennedy (1993), the higher cognitive load needed for completing the more complex task in our setting makes it likely that a higher proportion of auditors refrain from considering all cues simultaneously, resorting instead to step-by-step information processing that is prone to recency. The occurrence of the recency effect in our study enables us to investigate potential mitigating effects of trait skepticism.

4, Panel B) and AMOS (Figure 2), with both approaches yielding basically the same results. We measure cognitive effort (*Effort*) by adding up the responses to the following two items of the post-experimental questionnaire: “I thought intensively about the evaluation of the entity’s ability to continue as a going concern” and “I was very concentrated while assessing the entity’s ability to continue as a going concern” (scale from 1 “totally disagree” to 7 “totally agree”).¹⁸

The path model includes the direct effect of trait skepticism (*PS*) on belief revision (*Belief_Revision*) and its indirect effect through cognitive effort (*Effort*). It uses a multi-group design where the manipulated order condition constitutes the two groups. Constraining/unconstraining one path at a time reveals that the effects of *PS* on *Belief_Revision* and of *PS* on *Effort* do not differ significantly across the two groups. However, the path from *Effort* to *Belief_Revision* differs marginally significantly across the two groups as will be discussed in more detail below. Accordingly, we specify the model by constraining the former paths and unconstraining the latter path. The resulting model has an acceptable fit (Chi-square = 1.26, df = 2, p-value = 0.532; see Figure 2 for further information on model fit).

Figure 2 and Table 4, Panel B show the results. First, we find a positive direct effect between trait skepticism (*PS*) and belief revision (*Belief_Revision*) (Coef. = 0.375, p-value = 0.013). However, because this effect is not significantly moderated by information order, it only indicates a level effect and does not suggest that trait skepticism mitigates the recency effect directly.

¹⁸ Self-reported effort has also been used in prior auditing research (e.g., Payne and Ramsay 2005). Our measure of self-reported effort is highly significantly related to the self-reported time needed for completing the study ($\rho = 0.352$, p-value < 0.001). As we explain in the development of the hypotheses, cognitive effort is only expected to play a role in the cognitively demanding setting where information is presented simultaneously. Accordingly, we do not find that effort plays a role in the first experiment where information is presented sequentially (neither as a main effect nor in interaction with order or trait skepticism).

Second, we investigate the indirect path of trait skepticism on belief revisions through effort. For the first link of trait skepticism (*PS*) on cognitive effort (*Effort*), we find a highly significant effect across both order conditions (Coef. = 0.08, p-value = 0.001, see also Table 4, Panel B, Model I). Splitting up trait skepticism into its subconstructs shows a highly significant effect of its evidence-related subconstruct *PS_Evid* (Coef. = 0.10, p-value = 0.004) but no significant effect for its other subconstructs *PS_Other* (see Table 4, Panel B, Model II).¹⁹ For the second link of *Effort* on *Belief_Revision*, we observe an interaction effect with information order. Comparing the model where this path is constrained (Chi-square of 4.1, df = 3) to the model where it is unconstrained (Chi-square = 1.26, df = 2) shows a marginally significant effect with a p-value of 0.092 (Δ Chi-square = 2.84, Δ df = 1). In particular, effort reduces the effects of recency when positive evidence is presented last (Coef. = -1.62, p-value = 0.080). These findings correspond to the negative and marginally significant interaction effect of *Effort* and *Order* in the regression model (see Table 4, Panel B, Model III) and provide evidence that cognitive effort mitigates the recency bias. Results remain inferentially similar when trait skepticism (*PS*) is split up into its evidence-related subconstruct (*PS_Evid*) and its other subconstructs (*PS_Other*) (see Table 4, Panel B, Model II and IV).

Finally, we explore whether other demographic variables, especially auditor experience, have a debiasing effect. However, we do not observe significant effects for these variables. Our results reported above remain inferentially similar when including those variables into our model.

¹⁹ Exploring the relation between the individual dimensions of trait skepticism and cognitive effort reveals significantly positive effects for all evidence-related dimensions and also for autonomy (p-value < 0.05), but no significant effects for the dimensions of interpersonal understanding and self-esteem.

In summary, the results for the path model provide evidence for an indirect effect of trait skepticism on recency through effort.

--- Insert Figure 2 here ---

--- Insert Table 4 here ---

6. Conclusions

Professional skepticism is a concept of utmost importance in auditing. The results of this study demonstrate that trait professional skepticism plays an essential role in auditors' information processing and cognitive performance. We find that trait skepticism, in particular its evidence-related subconstruct, mitigates the recency effect. When information is presented step-by-step, auditors scoring higher on the evidence-related dimensions of trait skepticism are less likely to overreact to contrasting pieces of evidence. When information is presented simultaneously and auditors have the possibility to integrate all pieces of evidence at the end of the sequence, we predict and find that trait skepticism mitigates the recency bias via higher cognitive effort. We derive these findings using a non-fraud setting to mitigate potential confounding effects from state skepticism.

Our findings have several important implications for audit practice and future research. Following the recommendation of Hurtt (2010), we consider whether some subconstructs of trait skepticism are most useful in predicting auditor behavior and find that the evidence-related subconstruct is useful in predicting evidence evaluation of auditors. Moreover, our results suggest that trait skepticism has the potential to counteract and alleviate the recency bias. We investigate the judgment process in more detail and find that evaluate evidence more carefully as suggested by the evidence-related dimensions of trait skepticism can shield auditors from overreacting to

contrasting evidence when information is presented step-by-step. Moreover, we find that trait skepticism fosters cognitive effort, facilitating the handling of the cognitive load involved when integrating multiple pieces of evidence into the final belief and ultimately alleviating the recency bias. Overall, these findings provide evidence that trait professional skepticism manifests itself in auditor judgment. Our study extends prior studies by demonstrating the mitigating effects of an innate characteristic, namely trait skepticism, for a group of experienced auditors. This finding is important given the robustness and pervasiveness of the recency bias as well as the fact the behavioral effects of traits are usually hard to elicit. This finding suggests that auditing firms may benefit from developing a firm culture that incentivizes and rewards trait skepticism as suggested by the IAASB (2015). Our large sample of experienced auditors derived in a controlled environment also enables us to further validate the scale of Hurtt (2010).

The findings reported in this study should be considered in light of its limitations. External validity issues typical for experimental research also apply to our study and may potentially limit the generalizability of the results beyond other settings, persons, and tasks. One important limitation is that our study focuses on information processing and not on information search. As the study was conducted with auditors from one Big 4 audit firm, firm-specific effects may have affected the results. However, as the cognitive phenomena and judgment processes explored in the present study generally operate subconsciously, there are no grounds for suspecting them to be influenced by firm culture or other firm-specific features.

References

- Asare, S. K. 1992. The auditor's going-concern decision: Interaction of task variables and the sequential processing of evidence. *The Accounting Review* 67 (2): 379–93.
- Ashton, A. H., and R. H. Ashton. 1988. Sequential belief revision in auditing. *The Accounting Review* 63 (4): 623–41.
- Ashton, R. H., and J. Kennedy. 2002. Eliminating recency with self-review: The case of auditors' "going concern" judgments. *Journal of Behavioral Decision Making* 15 (3): 221–31.
- Bamber, E. M. M., R. J. Ramsay, and R. M. Tubbs. 1997. An examination of the descriptive validity of the belief-adjustment model and alternative attitudes to evidence in auditing. *Accounting, Organizations and Society* 22 (3/4): 249–68.
- Behling, O., and K. S. Law. 2000. *Translating questionnaires and other research instruments: Problems and solutions*. London, UK: Sage Publications.
- Bhaskar, L. S., T. Majors, and A. Vitalis. 2015. *Are "good" auditors impacted more by depletion? Threats to valued auditor attributes*. Working paper, Indiana University, University of Illinois at Urbana-Champaign, and Georgia Institute of Technology.
- Bonner, S. E. 1990. Experience effects in auditing: The role of task-specific knowledge. *The Accounting Review* 65 (1): 72–92.
- Butt, J. T., and T. I. Campbell. 1989. The effects of information order and hypothesis-testing strategies on auditors' judgments. *Accounting, Organizations and Society* 14 (5/6): 471–79.
- Carpenter, T. D., and J. L. Reimers. 2013. Professional skepticism: The effects of a partner's influence and the presence of fraud on auditors' fraud judgments and actions. *Behavioral Research in Accounting* 25 (2): 45–69.

Cushing, B. E., and S. S. Ahlawat. 1996. Mitigation of recency bias in audit judgment: The effect of documentation. *Auditing: A Journal of Practice and Theory* 15 (2): 110–22.

Cushing, B. E. 2000. Economic analysis of skepticism in an audit setting. *14th Symposium on Auditing Research*. University of Illinois at Urbana-Champaign, 1–3.

Doucet, M. S., and T. A. Doucet. 1996. Ethics, experience and professional skepticism: A situational analysis. *Behavioral Research in Accounting* 8 (Supplement): 158–68.

Fahrer, M. 2012: Gathering the evidence. *Charter: Journal of the Institute of Chartered Accountants in Australia* 83 (8): 30–32.

Fiske, S. T. 1980. Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology* 38 (6): 889–906.

Gibbins, M. 1984. Propositions about the psychology of professional judgement in public accounting. *Journal of Accounting Research* 22 (1): 103–25.

Griffith, E. E., J. S. Hammersley, K. Kadous, and D. Young (2015). Auditor Mindsets and Audits of Complex Estimates. *Journal of Accounting Research* 53 (1), 49-77.

Gudmundsson, E. 2009. Guidelines for translating and adapting psychological instruments. *Nordic Psychology* 61 (2): 29-45.

Hogarth, R. M., and H. J. Einhorn. 1992. Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology* 24 (1): 1–55.

Hurt, K., M. Eining, and R. D. Plumlee. 2008. *An experimental examination of professional skepticism*. Working paper, Baylor University and University of Utah.

Hurt, R. K. 2010. Development of a scale to measure professional skepticism. *Auditing: A Journal of Practice and Theory* 29 (1): 149–71.

Hurt, R. K., H. Brown-Libur, C. E. Earley, and G. Krishnamoorthy. 2013. Research on auditor professional skepticism: Literature synthesis and opportunities for future research. *Auditing: A Journal of Practice and Theory* 32 (Supplement 1): 45–97.

IAASB. 2012. Professional skepticism in an audit of financial statements: Staff questions and answers. <http://www.ifac.org/publications-resources/staff-questions-answers-professional-skepticism-audit-financial-statements>.

IAASB. 2014. Work plan for 2015 – 2016: Enhancing audit quality and preparing for the future. <https://www.ifac.org/system/files/publications/files/IAASB-Work-Plan-2015-2016.pdf>.

IAASB. 2015. Enhancing audit quality in the public interest: A focus on professional skepticism, quality control and group audits. <http://www.ifac.org/system/files/publications/files/IAASB-Invitation-to-Comment-Enhancing-Audit-Quality.pdf>.

Kahle, J., R. Pinsker, and R. Pennington. 2005. Belief revision in accounting: A literature review of the belief-adjustment model. *Journal of Accounting Literature* 8: 1–40.

Kennedy, J. 1993. Debiasing audit judgment with accountability: A framework and experimental results. *Journal of Accounting Research* 31 (2): 231–45.

Knechel, W. R., and W. F. Messier. 1990. Sequential auditor decision making: Information search and evidence evaluation. *Contemporary Accounting Research* 6 (2): 386–406.

Krishnamoorthy, G., T. J. Mock, and M. T. Washington. 1999. A comparative evaluation of belief revision models in auditing. *Auditing: A Journal of Practice and Theory* 18 (2): 105–27.

Krull Jr, G., P. M. J. Reckers, and B. Wong-on-Wing. 1993. The effect of experience, fraudulent signals and information presentation order on auditors' beliefs. *Auditing: A Journal of Practice and Theory* 12 (2): 143–53.

Lagnado, D. 2011. Thinking about evidence. *Proceedings of the British Academy* 171: 183–223.

Maurer, T. J., and R. A. Alexander. 1991. Contrast effects in behavioral measurement: An investigation of alternative process. *Journal of Applied Psychology* 1 (76): 3–10.

McCrae, R. R., and P. T. Costa Jr. 1996. Toward a new generation of personality theories: Theoretical contexts for the five-factor model. *The five-factor model of personality: Theoretical perspectives*, ed. J. S. Wiggins, 51-87. New York, London: The Guilford Press.

McMillan, J. J., and R. A. White. 1993. Auditors' belief revisions and evidence search: The effect of hypothesis frame, confirmation bias, and professional skepticism. *The Accounting Review* 68 (3): 443–65.

Messier Jr, W. F., and R. M. Tubbs. 1994. Recency effects in belief revision: The impact of audit experience and the review process. *Auditing: A Journal of Practice and Theory* 13 (1): 57–72.

Nelson, M. W. 2009. A model and literature review of professional skepticism in auditing. *Auditing: A Journal of Practice and Theory* 28 (2): 1–34.

Payne, E. A., and R. J. Ramsay. 2005: Fraud Risk Assessments and Auditors' Professional Skepticism. *Managerial Auditing Journal* 20 (3): 321–30.

PCAOB. 2012. Staff audit practice alert no. 10: Maintaining and applying professional skepticism in audits. http://pcaobus.org/standards/qanda/12-04-2012_sapa_10.pdf.

Pennington, N., and R. Hastie. 1986. Evidence evaluation in complex decision making. *Journal of Personality and Social Psychology* 51 (2): 242–58.

Peytcheva, M. 2014. Professional skepticism and auditor cognitive performance in a hypothesis-testing task. *Managerial Auditing Journal* 29 (1): 27–49.

Popova, V. 2013. Exploration of skepticism, client-specific experiences, and audit judgments. *Managerial Auditing Journal* 28 (2): 140–60.

Quadackers, L., T. Groot, and A. Wright. 2014. Auditors' professional skepticism: Neutrality versus presumptive doubt. *Contemporary Accounting Research* 31 (16): 639–57.

Rasso, J. T. 2015. Construal instructions and professional skepticism in evaluating complex estimates. *Accounting, Organization and Society* 46 (Special issue): 44–55.

Robinson, S. N., M. B. Curtis, and J. Robertson. 2013. *A person-situation approach to the examination of professional skepticism: Consideration of time pressure and goal framing*. Working paper, Sam Houston State University and University of North Texas.

Shaub, M. K., and J. E. Lawrence. 1996. Ethics, experience and professional skepticism: A situational analysis. *Behavioral Research in Accounting* 8 (Supplement): 124–57.

Strack, F., Erber, R., and Wicklund R. A. (1982). Effects of salience and time pressure on ratings of social causality. *Journal of Experimental Social Psychology*, 18(6), 581–94.

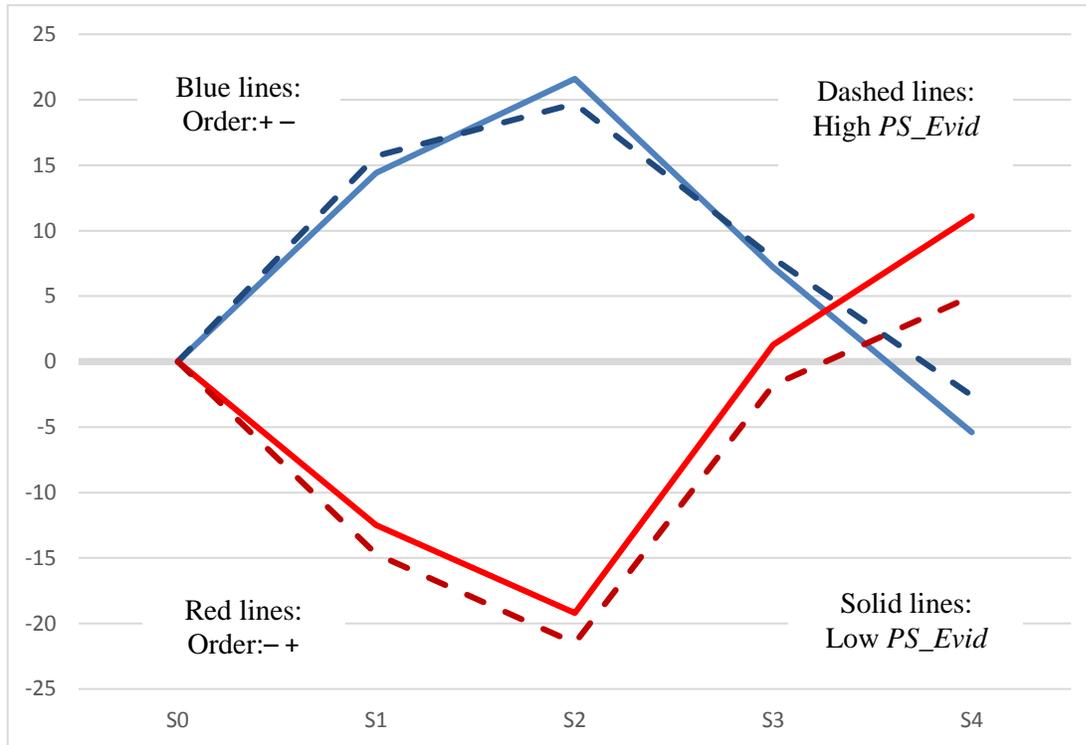
Trotman, K. T. 1998. Audit judgment research - Issues addressed, research methods and future directions. *Accounting and Finance* 38 (2): 115–56.

Trotman, K. T., and A. Wright. 1996. Recency effects: Task complexity, decision mode, and task-specific experience. *Behavioral Research in Accounting* 8: 175–93.

Trotman, K. T., and A. Wright. 2000. Order effects and recency: Where do we go from here? *Accounting and Finance* 40 (2): 169–82.

Tubbs, R. M., W. F. Messier Jr, and W. R. Knechel. 1990. Recency effects in the auditor's belief-revision process. *The Accounting Review* 65 (2): 452–60.

Figure 1
Experiment 1: Belief adjustment in a step-by-step response mode

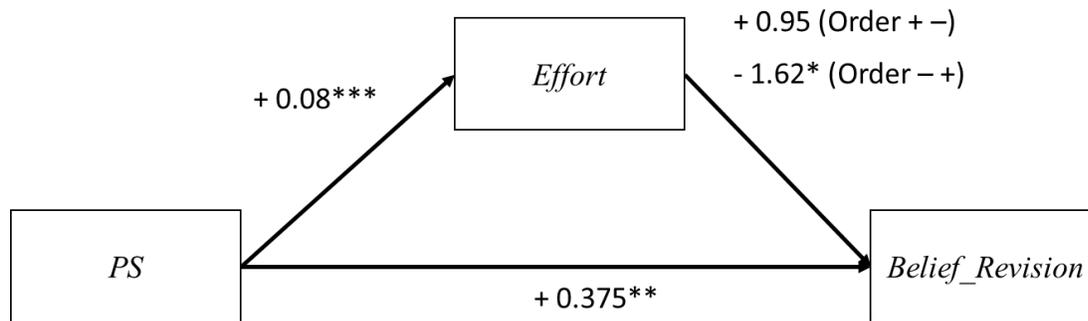


The figure shows the adjustment of the belief that the company will continue as a going concern after the presentation of each of the four additional pieces of evidence (S1-S4) relative to the first belief (S0). Auditors formed their first belief based on the background information provided.

The figure shows that auditors receiving positive information last (red lines) have, overall, a more positive belief revision than the auditors receiving negative information last (blue lines). Further, it shows that this general tendency is less pronounced for auditors with an above-median score on evidence-related professional skepticism (high *PS_Evid*; dashed lines) than for auditors with a below-median score (low *PS_Evid*; solid lines). Formal tests of these conclusions are reported in the result section.

Figure 2

Experiment 2: Belief adjustment in an end-of-sequence response mode



We employ a multi-group design where the order conditions constitute the two groups. The unconstrained model is not identifiable ($df = 0$). The constrained model has a Chi-square of 4.10 ($df = 3$). Unconstraining one path at a time reveals no significant differences between both groups for the effects of *PS* on *Effort* and for *PS* on *Belief_Revision*, but a marginally significant effect for the effect of *Effort* on *Belief_Revision*. The resulting model has a Chi-square of 1.26 ($df = 2$). The difference between this model and the fully constrained model is $df = 1$ and Chi-square = 2.84, indicating a p-value of 0.092.

The model fit is acceptable. The null hypothesis that there is no better model is not rejected (Chi-square = 1.26, $df = 2$, p-value = 0.532). The adjusted goodness of fit (AGFI) is 0.945, the comparative fit index (CFI) is 1.00 and the incremental fit index (IFI) is 1.029.

PS: Trait skepticism as measured by the Hurtt scale.

Effort: Self-rated intensity of thought and level of concentration in the performance of the experimental task.

Belief_Revision: Difference between final (S8) and initial belief (S0).

Table 1
Characteristics of participants

Panel A: Experience

Experience measure	Experiment 1 (n = 161)				Experiment 2 (n = 92)			
	n	Mean	SD	Range	n	Mean	SD	Range
Age [years]	138	36.9	8.2	26-59	88	39.4	8.1	27-59
Audit experience [years]	158	10.9	7.6	3-32	90	13.1	7.3	4-30
Industry experience [1-7]	161	2.5	1.5	1-7	92	2.0	1.3	1-6
GC experience [1-7]	161	3.9	1.3	2-6	92	4.4	1.4	2-7

Panel B: Hierarchical rank

Hierarchical rank	Experiment 1 (n = 161)		Experiment 2 (n = 92)	
	n	Proportion	n	Proportion
Assistant/senior	69	42.8%	18	19.6%
Manager/senior manager	64	39.8%	56	60.8%
Partner/director	25	15.5%	17	18.5%
Missings	3	1.9%	1	1.1%

Table 2
Professional skepticism

Professional skepticism and its subconstructs	Assistant/ senior (n = 87)	Manager/ senior manager (n = 120)	Partner/ director (n = 42)	Overall (n = 253)
Total score	133.3 (14.3)	135.5 (14.6)	137.8 (13.8)	135.3 (14.3)
Evidence-related				
Suspension of judgment	21.4 (3.4)	21.2 (4.2)	20.5 (5.2)	21.2 (4.1)
Search for knowledge	23.3 (4.0)	23.7 (3.7)	23.6 (4.1)	23.6 (3.9)
Questioning mind	11.7 (2.5)	11.5 (2.5)	11.6 (2.6)	11.6 (2.5)
Human-related				
Interpersonal understanding	22.7 (4.5)	22.7 (4.5)	22.9 (4.3)	23.0 (4.2)
Action-related				
Self-esteem	23.1 (3.7)	24.5 (3.2)	25.6 (3.0)	24.2 (3.4)
Autonomy	22.5 (3.5)	22.9 (3.7)	24.3 (3.0)	23.0 (3.6)

Reported are mean values (standard deviation in brackets).

The column “overall” also includes four participants with missing data on hierarchical rank.

Two items of the questioning mind category do not load on that construct (as in Hurtt 2010). We exclude them from the calculation of the subcomponent, but include them in the calculation of the total score.

Table 3**Experiment 1: Sequential information presentation****Panel A: Assessed likelihood that the company will continue to exist [in percent] by evidence-related trait skepticism score**

Step	Information order: +-			Information order: -+		
	High <i>PS_Evid</i> (n = 39)	Low <i>PS_Evid</i> (n = 39)	Overall (n = 78)	High <i>PS_Evid</i> (n = 38)	Low <i>PS_Evid</i> (n = 45)	Overall (n = 83)
S0	49.0	56.4	52.7	55.5	54.7	55.0
S1	64.7	70.8	67.8	40.8	42.2	41.6
S2	68.7	78.0	73.3	34.0	35.5	34.8
S3	56.9	63.6	60.3	53.7	56.0	54.9
S4	46.4	51.0	48.7	60.5	65.8	63.4
S2-S0	+19.7	+21.6	+20.6	-21.5	-19.2	-20.2
S4-S2	-22.3	-27.0	-24.6	+26.5	+30.3	+28.6
S4-S0	-2.6	-5.4	-4.0	+5.0	+11.1	+8.4

Reported are mean values. Standard deviations are omitted for the sake of brevity.

PS_Evid: Aggregate score of all evidence-related dimensions of the Hurtt scale: suspension of judgment, search for knowledge, questioning mind. For illustrating group effects, a median split is performed (high/low *PS_Evid*).

Panel B: Regression analysis

<i>Dependent variable: S4-S0</i>	(I) Coef./t-stat	(II) Coef./t-stat	(III) Coef./t-stat
<i>Order</i>	12.69*** (4.77)	12.65*** (4.75)	12.50*** (4.72)
<i>PS</i>	0.13 (1.37)		
<i>Order * PS</i>	-0.24 (-1.33)		
<i>PS_Evid</i>		0.07 (0.36)	0.01 (0.03)
<i>Order * PS_Evid</i>		-0.68* (-1.72)	-0.75* (-1.87)
<i>PS_Other</i>		0.17 (0.96)	0.23 (1.32)
<i>Order * PS_Other</i>		0.07 (0.22)	0.04 (0.12)
<i>ExpInd</i>			0.43 (0.48)
<i>Order * ExpInd</i>			-3.96** (-2.19)
<i>Adj. R²</i>	0.1195	0.118	0.134
<i>N</i>	161	161	161

Reported are coefficients (t-stats in brackets). ***(**)[*]: significant at the 1 percent (5 percent)[10 percent]-level (two-tailed). Model includes intercept (untabulated). All continuous variables are mean-centered.

Order: indicator variable; 1 if positive evidence is presented last; 0 otherwise.

PS: Trait skepticism as measured by the Hurtt scale.

PS_Evid: Aggregate score of all evidence-related dimensions of the Hurtt scale: suspension of judgment, search for knowledge, questioning mind.

PS_Other: Aggregate score of the human-related and action-related dimensions of the Hurtt scale: interpersonal understanding, self-esteem, autonomy.

ExpInd: Self-rated experience in the industry of the client (scale from 1 to 7).

Table 4
Experiment 2: Simultaneous information presentation

Panel A: Assessed likelihood that the company will continue to exist [in percent] by self-rated cognitive effort

Step	Information order: +-			Information order: -+		
	High Effort (n = 23)	Low Effort (n = 20)	Overall (n = 43)	High Effort (n = 19)	Low Effort (n = 30)	Overall (n = 49)
S0	59.1	60.5	59.8	58.5	63.7	61.6
S8	52.6	44.5	48.8	55.3	67.0	62.4
S8-S0	-6.5	-16.0	-10.9	-3.2	+3.3	+0.8

Reported are mean values. Standard deviations are omitted for the sake of brevity.

Effort: Self-rated intensity of thought and level of concentration in the performance of the experimental task. For illustrating group effects, we performed a median split (high/low *Effort*).

Panel B: Regression analysis

	(I) DV: <i>Effort</i> Coef./t-stat	(II) DV: <i>Effort</i> Coef./t-stat	(III) DV: <i>Belief_Revision</i> Coef./t-stat	(IV) DV: <i>Belief_Revision</i> Coef./t-stat
<i>PS</i>	0.08*** (4.20)		0.37** (2.39)	
<i>PS_Evid</i>		0.10*** (2.92)		0.45* (1.70)
<i>PS_Other</i>		0.62 (1.64)		0.30 (1.04)
<i>Order</i>			15.99** (-2.37)	16.22** (-2.39)
<i>Effort</i>			1.61* (-1.66)	1.67* (-1.71)
<i>Order * Effort</i>			-2.57* (-1.70)	-2.63* (-1.73)
<i>Adj. R²</i>	0.154	0.155	0.137	0.129
<i>N</i>	92	92	92	92

Panel B shows the results for a regression-based analysis of the path model (see Figure 2, for the path model using AMOS). Reported are coefficients (t-stats in brackets). All continuous variables are mean-centered.

***(**)[*]: significant at the 1 percent (5 percent)[10 percent]-level (two-tailed). Model includes intercept.

PS: Trait skepticism as measured by the Hurtt scale.

PS_Evid: Aggregate score of all evidence-related dimensions of the Hurtt scale: suspension of judgment, search for knowledge, questioning mind.

PS_Other: Aggregate score of the human-related and action-related dimensions of the Hurtt scale: interpersonal understanding, self-esteem, autonomy.

Effort: Self-rated intensity of thought and level of concentration in the performance of the experimental task.

Belief_Revision: Difference between final (S8) and initial belief (S0).