Ego depletion and auditors' JDM quality

Patrick J. Hurley

D'Amore-McKim School of Business, Northeastern University, 360 Huntington Ave, 404G Hayden Hall, Boston, MA, 02115, USA

ARTICLE INFO

Article history:
Received 10 November 2017
Received in revised form
26 February 2019
Accepted 1 March 2019
Available online xxx

Keywords:
Ego depletion
Self-control
Audit quality
Judgment and decision-making
Hypothesis generation
Cognitive processing

ABSTRACT

I report two experiments with senior-associate auditors that investigate whether self-control requirements in auditing tasks cause ego depletion, and whether depleting tasks impact auditors' judgment and decision-making (JDM). In Experiment 1, I find that using self-control to maintain vigilance/focus leads to greater levels of depletion than does using self-control to engage in cognitive processing in an audit planning task or inhibit impulses in a task from the psychology literature. However, I do not find that task-specific experience significantly reduces self-control resources required for task performance. In Experiment 2, I find evidence that depleting tasks — compared to a non-depleted control group — significantly reduce auditors' cognitive processing in the form of auditors' ability to generate plausible alternative hypotheses for client-provided explanations for trends. Finally, I find that depleting tasks reduce auditors' confidence in task performance, when compared to a non-depleted control group.

© 2019 Elsevier Ltd. All rights reserved.

1. Introduction

Recent research has stressed individual auditors' JDM as a significant factor for improving overall audit quality (Knechel, Krishnan, Pevzner, Shefchik, & Velury, 2013). Ego depletion (hereafter also "depletion") represents a causal mechanism that can potentially impact individual auditors' JDM quality and thereby audit quality. Ego depletion is defined as a temporary state in which prior self-control use reduces an individual's self-control resources. This reduction decreases willingness to engage in subsequent self-control acts (Baumeister, Bratslavsky, Muraven, & Tice, 1998; Baumeister & Vohs, 2016a). This paper reports the results of two experiments that investigate whether auditing tasks lead to ego depletion and the impact of these depleting tasks on auditors' judgment and decision-making (JDM) quality. The experiments answer calls to study potential causal mechanisms that impact JDM quality through psychological theory (Griffith, Kadous, & Young, 2016) and determine whether depletion is only caused by impulse inhibition or whether other self-control requirements cause depletion (Baumeister & Vohs, 2016a).

Self-control, the ability to control one's own behavior to conform to standards or pursue long-term goals (Baumeister, Vohs, & Tice, 2007), is crucial in many of auditors' JDM processes. For example, reasoning, cognitive processing, maintaining focus, and being vigilant all require self-control. As a result, self-control can be a determinant of auditors' JDM quality (Hurley, 2015) and it is important to study whether depleting tasks significantly impact auditors' JDM quality.

The current study addresses two main questions. First, does the type, rather than simply the presence, of a task's self-control requirements impact the amount of depletion from task performance, and does depletion result from applications of self-control other than simply inhibiting automatic responses (Baumeister & Vohs, 2016a)? Second, do depleting tasks impact auditors' cognitive processing? Experiment 1 addresses the former question, while Experiment 2 addresses the latter. It is unclear whether the causes and consequences of depletion and depleting tasks will extend to professional tasks because much of prior depletion research ignores important ecological features of the task environment by using mundane tasks that center on impulse inhibition. Indeed, updates to depletion theory suggest that inhibiting impulses, rather than other applications of self-control, drives the depletion effect (Baumeister & Vohs, 2016a). As a result, it is important to study whether tasks that go beyond simple inhibition lead to ego depletion. Further, prior research has largely ignored that task-specific experience can mitigate depletion by using tasks that are not familiar or meaningful to participants.

In Experiment 1, I investigate the causes of depletion using a 1 × 3 (self-control requirement: cognitive processing versus impulse inhibition versus maintaining vigilance/focus) between-subjects experiments using senior-associate auditors. Auditors in
Experiment 1 perform a task manipulated as primarily requiring one of these three types of self-control where the inhibition task serves as a benchmark because it leads to significant depletion in the psychology literature (e.g., Baumeister et al., 1998). I then measure participants’ depletion via their reaction time on an attention-control task (Stroop, 1935). This allows me to determine whether depletion varies based upon the type of self-control used on the initial task. Further, I am also able to examine whether task-specific experience reduces self-control resources required to perform a task.

I find that a task requiring auditors to maintain vigilance/focus leads to significantly greater levels of observed depletion than does the impulse inhibition task. An audit planning task, which requires cognitive processing, leads to depletion levels similar to those in the inhibition task. Further, I find that auditors’ task-specific experience does not significantly reduce the amount of self-control resources required for task performance. That is, greater task-specific experience does not significantly reduce the amount of depletion resulting from task performance.

In Experiment 2, I investigate the consequences of my depleting tasks using a 1 × 4 (self-control requirements from Experiment 1 and a control treatment) between-subjects experiment using seminatively trained auditors. Auditors in Experiment 2 complete the same manipulated task as in Experiment 1 and then complete a task in which they are required to provide as many plausible attention-control task (Stroop, 1935). This allows me to determine whether depletion varies based upon the type of self-control used on the initial task. Further, I am also able to examine whether task-specific experience reduces self-control resources required for task performance. That is, greater task-specific experience does not significantly reduce the amount of depletion resulting from task performance.

2 Despite moderate effect sizes reported in Hagger et al. (2010), some researchers have conducted additional meta-analyses that call into question the size of the ego depletion effect (Carter, Koller, Forster, & McCullough, 2015; Carter & McCullough, 2014). However, limitations in sampling procedures, methodological choices, and statistical analyses call into question the validity of these meta-analyses (cf. Cunningham & Baumeister, 2016). Various laboratories recently conducted a pre-registered replication and had difficulties replicating the depletion effect (Hagger et al., 2016); however, the standardized task implemented in this replication did not appropriately capture ego depletion (cf. Baumeister & Vohs, 2016b) and had significant differences in design from the study it intended to replicate (cf. Sprada, Kessler, & Ionides, 2016). On these bases, these challenges to the efficacy of the strength model of depletion should be interpreted with caution.

In additional analysis, I also find that auditors performing depleting tasks exhibit lower confidence in their hypothesis generation ability, and reduces their confidence in their task performance. This provides an understanding of the impact of depleting tasks on auditors’ JDM quality (Hurley, 2015) and contributes to a growing body of research that examines the effects of depleting tasks on auditors’ JDM quality (e.g., Bhaskar, Majors, & Vitalis, 2016; Mullis & Hatfield, 2018). Importantly, while Mullis and Hatfield (2018) report results consistent with depletion hindering auditors’ vigilance in finding seeded errors, I provide evidence regarding auditors’ cognitive processing, which is distinct and can have far-reaching implications for auditors’ JDM quality.

Finally, this research contributes to ego depletion theory and auditing practice by finding that task-specific experience does not significantly reduce depletion incurred by a task. This suggests that while self-control stamina can be increased with practice, as suggested in prior research, task-specific experience does not significantly reduce self-control demands from task performance.

The remainder of the paper proceeds as follows. In Section 2, I discuss ego depletion theory and related literature, and develop my hypotheses. In Sections 3 and 4 I discuss the participants, design and procedures, and results of Experiments 1 and 2, respectively. Section 5 offers conclusions, limitations, and suggestions for future research.

2. Related literature and hypothesis development

2.1. Ego depletion theory

Ego depletion theory is based upon the strength model of self-control (hereafter “the strength model”). The strength model analogizes using self-control, or willpower, to using a muscle in one’s body (Baumeister et al., 1998; Muraven & Baumeister, 2000). Specifically, Muraven, Tice, and Baumeister (1998) and others find that the willingness to exercise self-control becomes temporarily fatigued with use, requires a period of rest to recover and restore performance, and improves with long-term practice or training. The strength model therefore predicts that an individual’s subsequent self-control performance will suffer after performing an initial self-control act. Because most of the prior depletion literature supports the strength model (see Hagger, Wood, Stiff, & Chatzisarantis, 2010 for a meta-analysis), I use the strength model...
as the theoretical basis for my testable hypotheses later in this section.

The strength model of ego depletion is predicated on five assumptions (Muraven & Baumeister, 2000). First, self-control is required to make decisions and initiate or interrupt various types of behavior. Second, self-control relies upon a finite set of cognitive resources that can become depleted. Third, individuals likely differ in their overall amount of self-control resources. Fourth, an individual's current level of self-control resources dictates their success in performing acts of self-control. Fifth, using self-control expends these resources, which reduces the set of resources available for future self-control acts. Of specific interest are assumptions two, four, and five. The second assumption highlights the occurrence of ego depletion. The fourth assumption indicates that the timing of task performance matters, if a self-control task is performed before, rather than after, other self-control acts. The fifth and final assumption underlies the findings that self-control efforts degrade with increased depletion.

I focus my discussion on causes and consequences of ego depletion that apply to the current research. Prior psychology research finds that complex cognitive processing (Schmeichel, Vohs, & Baumeister, 2003), maintaining vigilance/focus (Vohs et al., 2008; Gailliot & Baumeister, 2005), and impulse inhibition (e.g., Baumeister et al., 1998) all potentially cause ego depletion. However, researchers have recently theorized that only impulse inhibition, as opposed to other applications of self-control, results in depletion (Baumeister & Vohs, 2016a). As a result, I center each of my manipulations on these three separate factors, which I discuss in greater detail in Section 3. Increased cognitive load (Schmeichel, 2007), as auditors encounter in multi-tasking settings (Mullis & Hatfield, 2018), and general decision-making (Bruyneel, Dewitte, Vohs, & Warlop, 2006; Gailliot & Baumeister, 2005; Vohs et al., 2008) also can potentially cause depletion and are inherent in many accounting tasks that auditors perform on a daily basis.

Prior research has also discovered a number of consequences stemming from ego depletion. Baumeister et al. (1998) find that depletion increases passivity and acceptance of the status quo. Importantly, the status quo in auditing is accepting the client's depletion increases passivity and acceptance of the status quo. Many accounting tasks that auditors perform on a daily basis. Prior research has also discovered a number of consequences stemming from ego depletion. Baumeister et al. (1998) find that depletion increases passivity and acceptance of the status quo. Importantly, the status quo in auditing is accepting the client's depletion increases passivity and acceptance of the status quo. Most depletion studies also use tasks that are not familiar or meaningful to task performers, thus largely ignoring the potential impact of task-specific experience on depletion. By contrast, auditors likely have experience with performing auditing tasks. This task-specific experience can potentially reduce the extent to which self-control resources are depleted by task performance by converting controlled and effortful processing into automatic and heuristic-based processing that is unaffected by depletion (Neal, Wood, & Drolet, 2013; Schmeichel et al., 2003). It is important to note that this effect is distinct from prior findings that practicing self-control exertion can increase self-control capacity (Oaten & Cheng, 2006a, 2006b, 2007; Muraven, 2010; Muraven, Baumeister, & Tice, 1999). Based on this discussion, it is possible that auditors are less susceptible to depletion from auditing tasks, compared to prior psychology studies' participants who are unlikely to have task-specific experience. To examine the effect of task-specific experience on depletion incurred from task performance, I posit the following hypothesis (in directional form):

H1. Different types of self-control use do not cause different levels of ego depletion.

Most depletion studies also use tasks that are not familiar or meaningful to task performers, thus largely ignoring the potential impact of task-specific experience on depletion. By contrast, auditors likely have experience with performing auditing tasks. This task-specific experience can potentially reduce the extent to which self-control resources are depleted by task performance by converting controlled and effortful processing into automatic and heuristic-based processing that is unaffected by depletion (Neal, Wood, & Drolet, 2013; Schmeichel et al., 2003). It is important to note that this effect is distinct from prior findings that practicing self-control exertion can increase self-control capacity (Oaten & Cheng, 2006a, 2006b, 2007; Muraven, 2010; Muraven, Baumeister, & Tice, 1999). Based on this discussion, it is possible that auditors are less susceptible to depletion from auditing tasks, compared to prior psychology studies' participants who are unlikely to have task-specific experience. To examine the effect of task-specific experience on depletion incurred from task performance, I posit the following hypothesis (in directional form):

H2. Task-specific experience reduces the extent to which self-control resources are depleted by task performance.

4 Hurley (2015) provides a more comprehensive discussion of the causes and consequences of ego depletion in an auditing environment.

5 Hagger et al. (2010) explore different types and spheres of self-control in their meta-analysis. However, I am referring to a lack of studies that purposefully manipulate the type of self-control within a single experiment. In short, I do not believe that a single meta-analysis provides the internal validity to rigorously investigate this issue and therefore do not believe the depletion literature sufficiently considers this issue.

6 I am specifically referring to the type of self-control exertion and not the intensity/magnitude of the exertion. I hold magnitude or length of self-control exertion constant by standardizing the time available to perform the self-control (manipulated) task.

Please cite this article as: Hurley, P. J., Ego depletion and auditors' JDM quality, Accounting, Organizations and Society, https://doi.org/10.1016/j.aos.2019.03.001
2.3. Consequences of ego depletion and accounting tasks (Experiment 2)

After investigating the causes of depletion, Experiment 2 examines whether depleting tasks affect auditors’ subsequent task performance. Prior research finds that depletion decreases controlled, effortful processing; decreases in controlled processing force individuals to rely more on automatic, heuristic processing (Masicampo & Baumeister, 2008; Pocheptsova, Amir, Dhar, & Baumeister, 2009; Schmeichel et al., 2003). Decreased controlled processing can potentially hinder overall evidence evaluation and pattern recognition (e.g., Hammersley, 2006) and the ability to generate alternative hypotheses for trends. Auditors’ hypothesis generation significantly impacts their subsequent performance in identifying actual errors (Heiman-Hoffman, Moser, & Joseph, 1995). Accordingly, depletion can reduce JDM quality through reduced hypothesis generation ability due to decreased controlled effortful processing. As a result, I posit the following hypothesis (in alternative form):

H3a. Depleting tasks will reduce auditors’ hypothesis generation ability.

Prior research also finds that depletion increases individuals’ passivity and the likelihood that they will accept the status quo (Baumeister et al., 1998). Further, prior literature finds that depletion increases confirmatory information processing and susceptibility to persuasion (Burkley, 2008; Fischer, Greitemeyer, & Frey, 2008). As a result, depletion can hinder auditors’ JDM quality by causing differences in attention to evidence (e.g., Hoffman, Joe, & Moser, 2003) or by relying on management's assertions as the status quo (e.g., Earley et al., 2008). These findings indicate that depletion will make auditors judge client-provided, fraudulent evidence because confidence in current task performance serves as an input into an individual’s decision to continue to perform that task. DeBono and Muraven (2013) find that depletion reduces individuals’ confidence in their predictions of future performance; however, prior research has not explored the impact of depleting tasks on individuals’ retrospective confidence subsequent to task completion. I therefore do not have an ex ante prediction regarding the impact of depleting tasks on retrospective confidence, and pose the following research question:

RQ1: Do depleting tasks significantly impact auditors’ retrospective confidence in their task performance?

3. Experiment 1

I design Experiment 1 to investigate whether different types of self-control requirements in an initial task lead to significantly different levels of ego depletion (H1). I also examine whether task-specific experience or expertise (H2) mitigates the incidence of depletion.

3.1. Design and procedures

Prior depletion research relies upon laboratory experiments to investigate the causes and consequences of depletion. This research uses a two-task paradigm, where participants perform an initial task manipulated as depleting or non-depleting and subsequently perform a task that measures depletion levels, typically through task persistence or the ability to suppress habitual responses. These methods provide valid measures of depletion because they require self-control to override a habitual response or urge to quit a task. Lower persistence or an inability to suppress habitual responses indicates greater depletion.

Before discussing the design and procedures for Experiment 1, it is useful to discuss my general experimental approach. I use two separate experiments in a two-task paradigm. Experiment 1 is designed to determine whether auditing tasks can lead to depletion using an established measure of depletion. Experiment 2 is designed to examine whether these exact same auditing tasks that cause depletion in Experiment 1 impact auditors’ JDM quality. I use this approach rather than conducting one experiment using a three-task paradigm, where I would deplete auditors, measure depletion, and subsequently assess depletion’s impact on JDM. 7

I made the design choice to run two separate experiments as opposed to one experiment using a three-task paradigm based upon theoretical concerns. 8 Prior literature provides evidence that two separate effects can influence the dependent variable in a depletion study using a three-task paradigm: conservation of self-control resources; and learned industriousness and adaptation. The conservation hypothesis within ego depletion theory posits that the more an individual is depleted, the more they are motivated to conserve self-control resources (e.g., Muraven, Shmueli, & Burkley, 2006; Tyler & Burns, 2009). As a result, if individuals know that they must perform an additional task that requires self-control beyond the current task, they will conserve self-control resources for the future task at the expense of the current task. This results in improved performance on the third task, based upon conservation of resources. However, Converse and DeShon (2009) provide evidence that a three-task paradigm leads to participants exhibiting learned industriousness and adaptation, which also increases performance on the third task. 9 Learned industriousness and adaptation therefore can confound depletion and self-control conservation in a three-task paradigm, making it difficult or impossible to disentangle which of these factors accounted for the observed results. 10 Further, attempting to isolate or control for conservation of self-control resources by explicitly discussing self-control requirements of tasks (as in Muraven et al., 2006) could potentially lead to demand or other uncontrolled effects (e.g., activation of beliefs in implicit theories of self-control; Job, Dweck, & Walton, 2010), which would weaken and/or confound my inferences and potentially introduce a significant amount of noise into my measures. As a result of these concerns, I determined that

---

7 This design choice is not without its drawbacks. I address these issues and alternative explanations for my reported results in subsection 4.4.11 and in the conclusion section.

8 There are also statistical concerns with this approach. Specifically, I would need to conduct a multilevel mediation model, as simple mediation analysis cannot handle clustered data from repeated measures (e.g., the Stroop task trials) and would yield biased standard errors of estimates (Krull & MacKinnon, 2001). However, multilevel mediation models and multilevel structural equation modeling cannot accommodate models with a dependent variable at the individual, as opposed to the trial, level (Krull & MacKinnon, 2001; Preacher, Zyphur, & Zhang, 2010, p. 211).

9 Converse and DeShon (2009, 1319) explicitly note that this phenomenon, learning tasks’ effort demands and transferring this knowledge and corresponding effort across similar tasks, is especially strong in situations involving more than two tasks. Further, because the task measuring depletion in my setting is dissimilar from the other two tasks, learned industriousness could lead to the same decreased performance on this task as would conservation of self-control resources.

10 For example, would high levels of performance on the second and third tasks indicate learned industriousness and adaptation, indicate that the participant was not depleted by the initial task, or indicate that the individual conserved self-control resources to improve performance on the third task? Using two separate experiments grounded in the two-task paradigm avoids this confound.
using two experiments was a more appropriate approach to test my hypotheses.

In Experiment 1, I use a 1 × 3 between-subjects design and randomly assign participants to one of three separate depletion treatments. I base my treatments upon an audit risk assessment task and label them as Processing, Inhibition, and Vigilance based upon the type of self-control required for task performance, which I describe in greater detail below. The Inhibition treatment has been used in prior literature and shown to induce depletion (Baumeister et al., 1998). As a result, the Inhibition treatment serves as a valid benchmark against which to compare the other treatments, and I use it as a quasi-control treatment rather than use a traditional control group. Not using a traditional control treatment biases against finding significant differences in depletion from my treatments, thus providing a stronger test. Subsequent to completing the manipulated risk assessment task, participants complete a Stroop color-word interference task (hereafter “Stroop task”). The Stroop task is designed to measure depletion levels through reaction times while inhibiting a habitual response. Finally, participants complete a post-experimental questionnaire.

Senior-associate auditors participate in Experiment 1 following training sessions. Auditors read a consent form, then begin the experimental task, which involves reading a client audit case. The risk assessment task is manipulated based upon the primary type of self-control required for task performance. Participants all work on this task for 15 min to ensure that the length of the depleting task is controlled across treatments.11 The Processing manipulation requires participants to use cognitive processing to complete risk assessment procedures from the planning phase of an audit. Specifically, participants perform a portion of the audit planning case used in Hammersley, Johnstone, and Kadous (2011), whereby they read information about a hypothetical client in order to identify risk factors and assess inherent, control, and fraud risk levels.12 I elect to use a risk assessment task because this task involves self-control through cognitive processing, which is present in many auditing tasks.13 Further, risk assessment tasks are common within audits, which increases the likelihood that auditors will have experience with this task. The Inhibition manipulation requires participants to read the case and cross out all instances of the letter “e.” Auditors complete the first page of the case by crossing out all instances of the letter “e” with no additional rules. This first page is designed to create a habitual behavior, which must be inhibited later in the task. Upon completing the first page, participants read updated instructions, which inform them that they must continue to cross out instances of the letter “e” according to several specific rules (e.g., do not cross out the letter “e” if it is immediately preceded or followed by a vowel). This task causes depletion because participants must override their habitual impulse of crossing out all instances of the letter “e” in order to apply these specific rules (Baumeister et al., 1998).

The Vigilance task requires participants to read case information and respond to a number of simple questions about the case. Participants are informed that the answers to these questions can be found chronologically throughout the case materials. This task therefore requires self-control in the form of maintaining vigilance/focus to scan through the case for the answers. Further, this task is similar to familiarizing oneself with a client prior to joining an audit team.

Upon completing the manipulated task, participants in Experiment 1 complete the Stroop task (Stroop, 1935) electronically via computer to measure depletion. The Stroop task is a psychological test of attention control, which requires participants to quickly identify the color of font in which words are displayed, rather than the actual meaning of the word. The Stroop task contains both congruent and incongruent trials. In congruent trials, the font color and meaning of the word are the same (e.g., “orange” presented in orange ink; orange is the correct response and is indicated by pressing the “o” key); in incongruent trials, the font color and meaning of the word differ (e.g., “orange” presented in green ink; green is the correct response and is indicated by pressing the “g” key). Incongruent trials require self-control in order to override the habitual impulse to respond to the meaning of the word, rather than the font color. Participants complete 40 trials of the Stroop task, balanced equally between congruent and incongruent trials. Trial order is randomly pre-drawn and held constant between participants to ensure greater experimental control.14 After completing the Stroop task, auditors complete a post-experimental questionnaire. Specifically, auditors provide demographic information, a self-reported measure of experience with tasks similar to the manipulated task (a 7-point Likert scale anchored on 1 = “No prior experience”; 4 = “Dealt with on a number of occasions”; and 7 = “Dealt with very often”), and a mood assessment.15 Auditors also complete trait measures of self-control (Brief Self-Control Measure; Tangney, Baumeister, & Boone, 2004) and professional skepticism (Hurtt, 2010).

3.2. Participants

Professional participants include 95 senior-associate auditors, who participated in Experiment 1 at the conclusion of training sessions. I removed seven auditors who did not complete the experimental materials, seven who did not appropriately complete the Stroop task, and four who used their phones to complete the Stroop task.16 Subsequent analyses and demographics are based
upon the remaining 77 auditors. Overall, 55.8 percent of auditors are male, and auditors’ mean (median) mood was 0.7 (0) on an 11-point Likert scale (anchored by: −5 = “Very unpleasant”; 5 = “Very pleasant”), indicating neutral mood on average. The mean (median) score on the Brief Self-Control Measure was 42.9 (44), and 126.9 (127.8) for trait Professional Skepticism. The number of auditors who completed the Processing, Inhibition, and Vigilance treatments was 23, 28, and 26 respectively.

3.3. Variables of interest

The dependent variable for the current analysis is participants’ reaction time, in seconds, on correctly answered incongruent Stroop task trials. I focus my analyses on incongruent Stroop trials only because of their self-control requirement that is not present in congruent trials. The main independent variable of interest for H1 is Treatment (Processing and Vigilance) in my repeated measures ANOVA (mixed-effects regression) below. I test H2 using a self-reported task-experience measure (T1Exp).

3.4. Results and discussion

3.4.1. Causes of ego depletion (H1 and H2)

In untabulated analyses, I find that auditors’ average reaction time for the Processing (1.57 s) and Inhibition (1.53 s) treatments are similar; however, average reaction time for the Vigilance treatment is 2.01 s. This initial comparison suggests that the Processing treatment causes similar levels of depletion as the Inhibition treatment found to cause significant depletion in prior psychology studies. However, the Vigilance treatment causes higher levels of depletion than did the other two treatments. This initial evidence contradicts the null hypothesis H1, which predicts that different types of self-control will not cause different levels of ego depletion. It is also important to note that these simple means do not account for other factors, like task-specific experience and learning.

To investigate H1, I first conduct a repeated-measures ANOVA with Time as the dependent variable and Treatment as the independent variable to assess whether my treatments significantly impact depletion. I include Question which is the overall number of the Stroop trial and controls for learning effects. I also include Participant as a control variable, as individuals may have different inherent reaction-time ability. Table 1, Panel A presents the results of this analysis. Treatment represents my test of H1, and is significant (F[2,1317] = 111.16, p < .001); this leads me to reject the null hypothesis H1 and indicates that my treatments cause different levels of depletion. Question (F[19,1317] = 14.08, p < .001) and Participant (F[76,1317] = 32.37, p < .001) are also significant, consistent with expectations.

To more fully investigate H1, and to explore H2, I use a mixed-effects linear regression. I use mixed-effects linear regression because, in addition to controlling for repeated measures, this model allows for easier interpretation of task experience and its directional relationship with the dependent variable that I use to test H2. Table 1, Panel B presents the following model and variables’ corresponding to H1 and H2:

\[
\text{Time} = \beta_0 + \beta_1 \text{Processing} + \beta_2 \text{Vigilance} + \beta_3 \text{Question} + \beta_4 \text{T1Exp} + \beta_5 \text{ProfSkep} + \epsilon
\]

Time is the reaction time for correct incongruent Stroop trials. Processing and Vigilance are binary variables that are set to 1 if the observation relates to a participant in the Processing or Vigilance treatment, respectively. Question is the overall question number of the Stroop trial to control for learning effects as the trials progress. T1Exp is the participant’s experience with tasks similar to the manipulated task and tests H2. Average experience with the manipulated task (T1Exp) is 3.54/1.37/3.96 for the Processing/Inhibition/Vigilance treatments, respectively. ProfSkep is the individual’s score on the professional skepticism measure. I specifically include ProfSkep due to the findings of Bhaskar et al. (2016), who find that professional skepticism can potentially exacerbate the effects of depletion.

Similar to my ANOVA analysis, and contrary to the null hypothesis H1, Vigilance is positive and significant (z = 2.47, two-tailed p = .013). Processing is positive but not significant at conventional levels (z = 0.90, two-tailed p > .35). This result expands our understanding of ego depletion theory by indicating that different types of self-control cause differences in depletion and that depletion is not only caused by impulse inhibition (Baumeister & Vohs, 2016a). Specifically, cognitive processing leads to similar levels of depletion as does an impulse inhibition task and maintaining vigilance/focus causes significantly greater depletion than either of these other forms of self-control. This finding also expands our understanding of factors affecting auditors’ JDM, given that realistic auditing tasks cause depletion and depletion can have significant consequences for auditors’ JDM (Hurley, 2015). H2 predicts that task-specific experience mitigates depletion. Inconsistent with H2, T1Exp is negative but is not significant at conventional levels (z = −1.00, one-tailed p > .15). This result indicates that task-specific experience does not reduce the self-control resources required to perform a task, which contradicts the findings of Neal et al. (2013).

Consistent with expectations, I find a significant learning effect as participants complete additional Stroop trials, as Question is negative and significant (z = −13.55, two-tailed p < .001). Trait professional skepticism (ProfSkep), however, is not significantly associated with depletion, either as a main effect or, in untabulated analyses, as an interaction with my treatments.

4. Experiment 2

I designed Experiment 2 to determine whether depleting tasks significantly affect JDM quality through hypothesis generation ability (H3a), susceptibility to fraudulent explanations (H3b), and confidence in task performance (RQ1). As discussed in section 3.1,
auditors manipulated depleting tasks into more meaningful measures of subsequent auditing JDM task. As a result, Experiment 2 links the association auditors as participants in both experiments to ensure the constant between Experiment 1 and Experiment 2 and use senior-found. Based upon these concerns, I hold the initial tasks con-
a three-task design would introduce potential theoretical con-
Flows. A CFO has given to explain troubling instead of the Stroop task. The auditing task involves evaluating the second task. In Experiment 2, auditors complete an auditing task are identical to those in Experiment 1 with the exception of the Control treatment. The design and procedures for Experiment 2 mirror those in Experiment 1 with the addition of a Vigilance (H1) +/− 0.557** (0.047) 0.047 (0.005) 0.003 (0.005) 0.003 0.003 0.003 0.003 0.003 0.003

Table 1
Participants’ reaction times – Stroop task.
Panel A: Repeated-Measures ANOVA (H1 Test)

<table>
<thead>
<tr>
<th>Source</th>
<th>Observations: 1,415</th>
<th>R²</th>
<th>0.689</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>924.85</td>
<td>.563</td>
<td>0.667</td>
</tr>
<tr>
<td>Treatment</td>
<td>70.40</td>
<td>2</td>
<td>35.20</td>
</tr>
<tr>
<td>Question</td>
<td>84.69</td>
<td>19</td>
<td>4.46</td>
</tr>
<tr>
<td>Participant</td>
<td>758.46</td>
<td>74</td>
<td>10.25</td>
</tr>
<tr>
<td>Residual</td>
<td>417.68</td>
<td>1,319</td>
<td>0.32</td>
</tr>
<tr>
<td>Total</td>
<td>1342.53</td>
<td>1,414</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Panel B: Mixed-Effects Linear Regression (H1 – H2 Test)

<table>
<thead>
<tr>
<th>Auditors</th>
<th>Pred.</th>
<th>Coeff (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>N/A</td>
<td>2.336***</td>
</tr>
<tr>
<td>Processing (H1)</td>
<td>+/-</td>
<td>0.194</td>
</tr>
<tr>
<td>Vigilance (H1)</td>
<td>+/-</td>
<td>0.857**</td>
</tr>
<tr>
<td>Question</td>
<td>–</td>
<td>–0.019***</td>
</tr>
<tr>
<td>T1Exp (H2)</td>
<td>–</td>
<td>–0.047</td>
</tr>
<tr>
<td>ProfSkep</td>
<td>+/-</td>
<td>–0.003</td>
</tr>
</tbody>
</table>

Time is the dependent variable and measures individuals’ response time (in seconds) to each incongruent, correct trial of the Stroop task. Time to complete the JDM task was held constant across all treatments. Proctors were asked specifically to watch for herding effects and did not report any herding effects. This design choice raises the possibility of alternative explanations for my observed results, which I discuss in detail in subsection 4.4.1.1.

4.1. Design and procedures

In Experiment 2, I use a 1 × 4 between-subjects design and randomly assign professional auditors to treatments. These treatments mirror those in Experiment 1 with the addition of a Control treatment. The design and procedures for Experiment 2 are identical to those in Experiment 1 with the exception of the second task. In Experiment 2, auditors complete an auditing task instead of the Stroop task. The auditing task involves evaluating explanations that a CFO has given to explain troubling financial trends of a company, generating plausible alternative explanations for those trends, and providing reasonableness ratings for the CFO’s explanations (Johnstone, Gramling, & Rittenberg, 2012, p. 670). The task is based upon the fraud at Koss Corporation, perpetrated by the CFO. In this case, the CFO-provided explanations are fraudulent. This task provides evidence of auditors’ cognitive processing measured by plausible alternative explanations and susceptibility to fraudulent explanations captured by reasonableness ratings. Auditors then rate their retrospective confidence in their performance and complete a post-experimental questionnaire, identical to that in Experiment 1. Auditors in the Control treatment complete the auditing task and post-experimental questionnaire.21

21 I elect not to use a filler task because my theoretical comparison of interest in using a Control treatment is to compare performance of depleted individuals to that of non-depleted individuals. The most reliable way to achieve a true no-depletion group is to simply measure performance on the task immediately. Importantly, Control treatment auditors had 30 min to complete their materials, while all depleting treatments had 45 min (15 min for the initial task plus the 30 min given to the Control group); this ensured that time to complete the JDM task was held constant across all treatments. Proctors were asked specifically to watch for herding effects and did not report any herding effects. This design choice raises the possibility of alternative explanations for my observed results, which I discuss in detail in subsection 4.4.1.1.
4.2. Participants

Professional participants include 127 senior-associate auditors, who participated in Experiment 2 at the conclusion of training sessions. Senior associates are appropriate participants for this task as they are well suited to the goals and task of the experiment (Libby, Bloomfeld, & Nelson, 2002, p. 802) and their decisions and documentation can influence decisions by their superiors (Ricchuite, 1999). I removed from analysis one auditor who did not complete the task, three auditors who did not complete the post-experimental questionnaire, and ten auditors who spent insufficient time on the tasks (i.e., less than 10 min combined). Subsequent analyses and demographics are therefore based upon 113 auditors. Overall, 51 percent of auditor participants were male, and their mean (median) mood was 0.7 (0) on an 11-point Likert scale (anchored by: −5 = “Very unpleasant”; 5 = “Very pleasant”), indicating a neutral mood on average. The mean (median) score on the Brief Self-Control Measure was 45.3 (46), and 131.2 (130) for trait Professional Skepticism. The number of auditors who completed the Processing, Inhibition, Vigilance, and Control treatments was 20, 21, 24, and 48 respectively.

4.3. Variables of interest

The dependent variables for my tests of depleting task consequences are: the number of valid (i.e., plausible) alternative explanations generated for each trend (H3a); participants’ reasonableness rating of the CFOs explanation for each trend (H3b); and participants’ self-reported retrospective confidence in task performance (RQ1). Two PhD students assessed and coded the validity of the provided explanations (H3a test). These individuals are former senior managers at a Big 4 accounting firm and are blind to experimental treatment during their coding. The coders met and resolved their coding issues. This reconciled coding is used in my following analyses.

4.4. Results and discussion

4.4.1. Generation of alternative explanations (H3a)

To examine H3a, I conduct analyses over ValidAlternatives. Table 2, Panel A provides the means for ValidAlternatives and RawAlternatives by treatment. Valid (Raw)Alternatives averaged 0.837 (1.204) for Processing, 0.942 (1.350) for Inhibition, 1.059 (1.395) for Vigilance, and 1.200 (1.417) for the Control treatment. I conduct an ANCOVA for ValidAlternatives with Treatment as the independent variable of interest. I also control for Trend and RawAlternatives, as valid hypotheses may differ depending on the specific trend and are necessarily a subset of the raw number of hypotheses generated. Table 2, Panel B provides the results of these tests.

H3a predicts that depleting tasks will reduce auditors’ generation of alternative explanations for the given trends. Consistent with H3a, Treatment is significant (F(3,551) = 5.23, p = .002); this indicates that differences in treatments led to significant differences in auditors’ ability to generate plausible alternative explanations for troubling financial trends. Trend, as expected, is significant (F(4,551) = 9.02, p < .001); this indicates that auditors were better able to generate plausible alternative explanations for certain trends than others. Finally, the number of RawAlternatives is significant in determining the number of ValidAlternatives (F(1,551) = 1039.76, p < .001); this is expected, as valid alternatives are a subset of raw alternatives, and controls for the total number of alternatives generated.

To further investigate the results of the ANCOVA, I perform a series of planned contrasts, which are presented in Table 3 Panel C. First, I compare ValidAlternatives for all depleting treatments to the Control treatment in order to compare the performance of depleting auditors to non-depleting auditors. I find that ValidAlternatives are significantly lower for depletion treatments than for the Control treatment (F(1,551) = 14.58, p < .001). Second, I test each treatment individually against the Control treatment and find that auditors in the Processing (F(1,551) = 9.24, p < .01), Vigilance (F(1,551) = 4.27, p < .05), and Inhibition (F(1,551) = 10.60, p < .01) treatments all generate significantly fewer valid alternative explanations for the trends. As a result, performing a depleting task, when compared to not being in a depleted state, significantly reduces auditors’ cognitive processing in the form of hypothesis generation.

Third, in order to test whether the statistically significant differences in depletion observed in Experiment 1 impact auditors’ JDM, I conduct three additional planned contrasts. Specifically, I compare ValidAlternatives in each depleting treatment to each of the two other depleting treatments. Contrary to expectations, I do not find significant differences in ValidAlternatives for the Vigilance treatment when compared to the Processing and Inhibition treatments, despite the Vigilance treatment causing significantly greater depletion in Experiment 1. Performance in the Vigilance treatment does not significantly differ from the Processing (F(1,551) = 0.96, p > .30) or Inhibition (F(1,551) = 1.28, p > .25) treatments. As expected, the Processing and Inhibition treatments do not significantly differ (F(1,551) = 0.02, p > .85).

4.4.1.1. Potential alternative explanations for H3a findings. Due to my experimental design, where I do not explicitly measure ego depletion in Experiment 2, it is possible that, while my results are theoretically consistent with ego depletion as an explanation, confounding factors other than ego depletion are driving my observed results. In this section, I identify potential alternative explanations for my observed results and discuss the likelihood that these alternative explanations are driving my findings.

I designed the control treatment to move directly to the JDM task, rather than simply deplete a auditor to mirror the depleting-task treatments. I chose this structure to ensure that my control treatment was truly a no-depletion treatment. While prior research has used tasks which ostensibly require little self-control (e.g., Muraven et al., 2006; Tice, Baumeister, Shmueli, & Muraven, 2007), self-control requirements exist on a continuum and the most reliable way to ensure the control group was not depleted was to forego an initial task. This approach raises the possibility that simple fatigue from completing an initial task led participants in the depleting-task treatments to generate fewer valid alternative explanations.

22 In untabulated analysis I added back these 13 auditors (14 exclusions less the auditor who did not complete the task) and re-ran my ANCOVA model for my test of H3a. My effect of Treatment in my ANCOVA is stronger with the inclusion of these observations (F(3,551) = 6.01 compared to F(3,551) = 5.23 reported) and all of my reported contrasts hold or improve their significance.

23 In untabulated analyses, I find the inter-rater reliability to be acceptable, based upon a Cohen’s Kappa of .757 for assessing the validity of alternative explanations (Peterson, 1994). For Trend 1 “Cash balances have declined to their lowest level since FY 2008,” examples of valid responses include: “Increased R&D expenditures,” “Paid off debt,” and “Net PP&E increased so the company is buying more long-term assets.” Examples of invalid responses include “Spending more cash on marketing” (inconsistent with SG&A numbers), “AR has increased due to decreased collectability” (inconsistent with A/R balance, which has not increased), and “COGS or other expenses have increased” (inconsistent with these balances).

24 All other collected variables (e.g., mood, trait self-control, trait professional skepticism, experience with tasks similar to the JDM task) are not significant when included in the model; therefore, I omit them for brevity and a lack of ex ante predictions regarding their effects.
Table 2
Participants’ alternative explanations (H3a Test).

<table>
<thead>
<tr>
<th>Condition</th>
<th>n</th>
<th>Mean Raw Alternatives (Std. Dev.)</th>
<th>Mean Valid Alternatives (Std. Dev.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processing</td>
<td>20</td>
<td>1.204 (1.093)</td>
<td>0.837 (0.992)</td>
</tr>
<tr>
<td>Inhibition</td>
<td>21</td>
<td>1.350 (1.045)</td>
<td>0.942 (0.978)</td>
</tr>
<tr>
<td>Vigilance</td>
<td>24</td>
<td>1.395 (1.051)</td>
<td>1.059 (0.985)</td>
</tr>
<tr>
<td>Control</td>
<td>48</td>
<td>1.417 (1.019)</td>
<td>1.200 (1.048)</td>
</tr>
</tbody>
</table>

Panel B: ANCOVA — Valid Alternative Explanations

<table>
<thead>
<tr>
<th>Source</th>
<th>Observations:</th>
<th>R²:</th>
<th>Root MSE:</th>
<th>Adj. R²:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>560</td>
<td>.7181</td>
<td>.545</td>
<td>.7140</td>
</tr>
<tr>
<td>Treatment (H3a)</td>
<td>545</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trend</td>
<td>551</td>
<td>175.43</td>
<td>8</td>
<td>9.02</td>
</tr>
<tr>
<td>RawAlternatives</td>
<td>551</td>
<td>1039.76</td>
<td>1</td>
<td>0.297</td>
</tr>
<tr>
<td>Residual</td>
<td>551</td>
<td>.297</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>559</td>
<td>1.039</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel C: Contrasts based on ANCOVA

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processing, Vigilance, and Inhibition &lt; Control</td>
<td>551</td>
<td>14.58</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Processing &lt; Control</td>
<td>551</td>
<td>9.24</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Vigilance &lt; Control</td>
<td>551</td>
<td>4.27</td>
<td>0.039</td>
</tr>
<tr>
<td>Inhibition &lt; Control</td>
<td>551</td>
<td>10.60</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Processing vs. Vigilance</td>
<td>551</td>
<td>0.96</td>
<td>0.328</td>
</tr>
<tr>
<td>Processing vs. Inhibition</td>
<td>551</td>
<td>0.02</td>
<td>0.898</td>
</tr>
<tr>
<td>Vigilance vs. Inhibition</td>
<td>551</td>
<td>1.28</td>
<td>0.259</td>
</tr>
</tbody>
</table>

Note: ValidAlternatives is the dependent variable in Panel B and measures the number of plausible alternative explanations generated. Treatment is a factor variable where 1, 2, 3, and 4 correspond to Processing, Inhibition, Vigilance, and Control treatments, respectively. Trend is a factor variable that controls for each of the five trends that participants evaluate. RawAlternatives is the total number of alternatives generated. All variables are presented with two-tailed p-values.

Table 3
Participants’ reasonableness ratings (H3b Test).

<table>
<thead>
<tr>
<th>Mean Rating (Std. Dev.)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Processing</td>
<td>4.158 (1.740)</td>
<td></td>
</tr>
<tr>
<td>Inhibition</td>
<td>4.233 (2.054)</td>
<td></td>
</tr>
<tr>
<td>Vigilance</td>
<td>4.319 (1.850)</td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>4.546 (1.768)</td>
<td></td>
</tr>
</tbody>
</table>

Panel B: ANOVA — Reasonableness Ratings

<table>
<thead>
<tr>
<th>Source</th>
<th>Observations:</th>
<th>R²:</th>
<th>Root MSE:</th>
<th>Adj. R²:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>560</td>
<td>.135</td>
<td>1.723</td>
<td>.122</td>
</tr>
<tr>
<td>Treatment (H3b)</td>
<td>559</td>
<td></td>
<td>31.778</td>
<td>10.71</td>
</tr>
<tr>
<td>Trend</td>
<td>551</td>
<td>3.573</td>
<td>20.614</td>
<td>18.44</td>
</tr>
<tr>
<td>TaskExp</td>
<td>551</td>
<td>2.968</td>
<td>2.968</td>
<td>0.009</td>
</tr>
<tr>
<td>Residual</td>
<td>559</td>
<td>3.380</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1,889.621</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Rating is the dependent variable, and measures participants’ reasonableness ratings for the CFO’s explanations of trends. Treatment is a factor variable where 1, 2, 3, and 4 correspond to Processing, Inhibition, Vigilance, and Control treatments, respectively. Trend is a factor variable that controls for each of the five specific trends that participants evaluate. TaskExp is participants’ self-reported experience with tasks similar to the JDM task. All variables are presented with two-tailed p-values.
for the observed trends. However, if fatigue were driving my results the prediction would be that participants in the depleting-task treatments generated fewer raw and valid explanations than did participants in the Control treatment. Depletion, by contrast, would not necessarily lead to fewer raw explanations, but would lead to fewer valid explanations through decreased complex cognitive processing.\textsuperscript{25} An untabulated ANCOVA indicates that Treatment does not significantly impact raw explanations ($F(1,356) = 1.01$, $p = .387$), but does significantly impact valid explanations, as reported previously. As a result, it does not appear that fatigue is driving my observed results.

A second possibility is that initial tasks requiring accounting knowledge (e.g., the Processing treatment) either 1) cause depletion through accessing this knowledge, or 2) prime accounting knowledge for use in the JDM task. In the former case, prior research finds that accessing knowledge does not lead to depletion, so it is unlikely that accessing accounting knowledge would be responsible for my findings (Baumeister & Vohs, 2016a). In the latter case, priming accounting knowledge on the initial task would lead to improved performance on the JDM task for auditors in the Processing treatment. By contrast, auditors in the Processing treatment exhibit the poorest performance among all treatments. As a result, it does not appear likely that accounting knowledge requirements are responsible for my observed results.

A third possibility is that auditors’ reactions to the initial task (e.g., interest in the task, enjoyment, etc.) influenced their performance on the JDM task in ways that did not occur for auditors in the Control treatment. While I did not measure auditors’ task interest or enjoyment, I measure their mood. Due to random assignment, differences in mood can be attributed to auditors’ affect resulting from the initial task manipulations or lack thereof in the Control treatment. In an untabulated ANOVA with Mood as the dependent variable and Treatment as the independent variable, I find that Treatment significantly influences auditors’ mood ($F(1,356) = 30.46$, $p < .001$, untabulated). However, I re-run my analyses of ValidAlternatives and include Mood as a direct-effect covariate and an interaction with Treatment and find that neither Mood nor the interaction term Treatment*Mood are significant (both $p > .75$, untabulated). Further, the effect of Treatment remains significant ($F(1,547) = 4.94$, $p = .002$, untabulated). As a result, it does not appear that auditors’ reactions to the initial manipulated task drive my observed results.

4.4.2. Susceptibility to persuasion (H3b)
Rating captures auditors’ rating, via seven-point Likert scale, of how likely they believe the CFO’s explanation accounts for each trend (anchored by: 1 = “Very unlikely”; 7 = “Very likely”). Table 3, Panel A provides the mean ratings by treatment for Experiment 2: 4.16 for Processing, 4.23 for Inhibition, 4.32 for Vigilance, and 4.55 for Control treatments. To investigate H3b, I conduct an ANOVA with Treatment as the independent variable of interest, and control for Trend, as in my test of H3a, and prior experience with tasks similar to the JDM task TaskExp. Table 3, Panel B provides the results of this analysis.

H3b predicts that performing depleting tasks will lead to higher reasonableness ratings for the CFO’s explanations. Inconsistent with H3b, Treatment is not significant at conventional levels ($F(3,551) = 1.20$, $p > .30$). This indicates that performing depleting tasks did not significantly impact auditors’ susceptibility to the CFO’s fraudulent explanations for trends. Trend, as expected, is significant ($F(4,551) = 18.44$, $p < .001$). TaskExp, as expected, is also significant ($F(1,551) = 6.95$, $p = .009$). I use an untabulated contrast to compare the depleting treatments to the Control treatment and find that depleting tasks do not significantly impact auditors’ reasonableness ratings ($F(1,551) = 1.92$, $p > .15$). Thus, performing depleting tasks does not significantly impact auditors’ susceptibility to fraudulent explanations in my setting.

4.4.3. Confidence (RQ1)
RQ1 considers whether depletion will significantly impact auditors’ confidence in task performance following the CFO explanation task. Auditors’ mean confidence by treatment is as follows: 4.23 for Processing, 4.57 for Inhibition, 4.75 for Vigilance, and 4.85 for the Control treatment. In an untabulated ANOVA, Treatment is not significant at conventional levels ($F(1,108) = 1.54$, $p > .20$), when controlling for TaskFamiliarity ($F(1,108) = 29.56$, $p < .001$). However, in an untabulated ANOVA comparing the Control treatment to all depletion treatments, auditors in the Control treatment reported higher mean confidence than depleted auditors ($F(1,108) = 9.72$, $p < .01$), indicating that depleting tasks reduce auditors’ retrospective confidence in task performance. This result complements that of DeBono and Muraven (2013), who find that depletion reduces individuals’ prospective confidence on a future task. Given that the depleting tasks hindered auditors’ ability to generate plausible alternative hypotheses, it is notable that these auditors then have lower confidence in their work because confidence in task performance is an input into the decision to continue to persist in performing a task. Indeed, this provides some indirect evidence that auditors are subconsciously aware of the impact of the depleting tasks on their performance.

5. Conclusions
I conduct two experiments to explore the causes and consequences of ego depletion with respect to professional auditing tasks. With respect to causes of depletion, I find that an auditing task requiring vigilance/focus causes greater levels of depletion than does a depleting psychology task based upon impulse inhibition. Further, an audit planning task requiring cognitive processing leads to similar levels of depletion as the depleting psychology task. These findings provide evidence that realistic, professional tasks can cause depletion at similar or even higher levels than a previously used psychological task of self-control, due to different self-control requirements. This finding contributes to theories of auditors’ JDM processes and ego depletion theory, which has recently theorized that depletion may stems solely from inhibiting impulses (Baumeister & Vohs, 2016a). This finding also underscores the PCAOB’s concerns regarding auditor workload; regulators should not only consider the amount of work, but whether specific types of auditing tasks have significant self-control requirements. I also find that task-specific experience did not significantly reduce the self-control resources required to perform a task.

With respect to consequences of ego depletion, I find evidence that performing depleting tasks, when compared to a non-depleted control group, significantly reduces auditors’ cognitive processing in the form of inhibiting their ability to generate plausible alternative hypotheses for troubling financial trends. However, auditors within the different depleting task treatments alone did not exhibit significantly different levels of JDM quality. This finding can inform
workpaper reviewers, as work performed after auditors have performed depleting tasks may require more thorough workpaper review. I also find that these depleting tasks reduce auditors’ retrospective confidence in task performance, which potentially guards against the effects of depletion if decreased confidence leads to increased effort. However, I do not find that depleting tasks lead to increased susceptibility to fraudulent explanations for these troubling trends. Importantly, auditors performed these tasks for less than 15 min, which would be considered a less-severe form of depletion. As a result, it is possible that more severe depletion (e.g., depletion occurring over multiple tasks, during multi-tasking, or during busy season) or more depleting tasks can more significantly impact auditors’ JDM processes. These findings also contribute to the accounting literature on ego depletion (e.g., Bhaskar et al., 2016; Hurley, 2015, 2017; Mullis & Hatfield, 2018).

The results of my study need to be interpreted with some caveats. First, for reasons outlined earlier, I did not use a three-task paradigm in my experimental design and choose to address my research questions using two separate experiments and to infer the effect of depletion from Experiment 1 to Experiment 2. I strengthen the tie between these experiments by using senior associate auditors in both experiments and using the exact same depleting tasks in both experiments. Second, in order to ensure that auditors in my control treatment were not depleted, they did not perform an initial task as did the auditors in my depleting-task treatments. While this design choice affords me greater control over a lack of depletion in the control treatment, it raises the possibility that confounding factors, rather than depletion, are driving my results. I discuss fatigue, differences in accounting knowledge requirements, and auditors’ reactions to the initial task as potential alternative explanations for my results and provide evidence and theoretical arguments to support that these factors do not drive my results.

Based upon my findings that depleting tasks hinder auditors’ complex cognitive processing but not their susceptibility to fraudulent explanations, future research should continue to investigate the consequences of ego depletion within professional accounting and auditing settings. For example, future research can examine whether more-severe levels of ego depletion, resulting from tasks that are either more complex or longer in duration, impairs JDM quality in predictable ways. Further, future research can investigate the impact of depletion on more complex tasks (e.g., audits of complex estimates), and whether existing mechanisms in the auditing environment (e.g., workpaper review and decision aids) can guard against the effects of depletion on auditors’ JDM.

Funding

This paper represents a portion of my dissertation and was funded by the University of Wisconsin - Madison Behavioral Research Insights Through Experiments (BRITE) Lab and Andersen Center. I thank my dissertation committee chair, Brian Mayhew, as well as my dissertation committee members Karla M. Zehms, Yuri Miyamoto, Justin Sydnor, and Terry Warfield for helpful guidance, discussions, and comments. I also thank anonymous accounting firms for providing professional auditors to participate in my experiments. Finally, I thank Jeff Cohen, Eric Condie, Emily Griffith, Jackie Hammersley, Kathryn Kadous, Matt Kaufman, Irina Malaescu, Kara Obermire, Nadra Pencle, Tim Rupert, Jason Smith (discussant), Amy Tegeler, Andrew Trotman, Ken Trotman (editor), anonymous reviewers, and participants at the University of Wisconsin – Madison, DePaul University, University of Texas at Austin, George Mason University, Northeastern University, and the 2017 AAA Auditing Midyear Meeting for helpful comments, assistance, and suggestions.

Data Availability

The laboratory data used in this study are available from the author upon request.

References


Please cite this article as: Hurley, P. J., Ego depletion and auditors’ JDM quality, Accounting, Organizations and Society, https://doi.org/10.1016/j.aos.2019.03.001.