Role of confidence in tax return preparation

Amy M. Hageman

Published Version Information

Citation: Hageman, A.M. (2010). The role of confidence in tax return preparation using tax software. In V. Arnold (Ed.), Advances in accounting behavioral research, volume 13 (pp.31-57). Bingley, UK : Emerald Group Publishing.

Digital Object Identifier (DOI): doi:10.1108/S1475-1488(2010)0000013006

Publisher's Link:
http://www.emeraldinsight.com/books.htm?chapterid=1870451&show=abstract

The Role of Confidence in Tax Return Preparation

Contact Information:

Amy M. Hageman
Research Fellow
Kenneth G. Dixon School of Accounting
University of Central Florida
4000 Central Florida Blvd.
Orlando, FL 32816-1400
Phone: 407-823-4420
FAX: 407-823-3881
E-mail: ahageman@bus.ucf.edu

THE ROLE OF CONFIDENCE IN TAX RETURN PREPARATION

Amy Hageman

University of Central Florida

ABSTRACT

This paper investigates the nature of tax preparers' confidence, as well as how the introduction of a tax decision support system (TDSS) affects tax preparers' confidence levels. Psychological theories of confidence (e.g., Einhorn & Hogarth, 1978) are drawn upon to develop predictions regarding the role of process (*ex-ante*) and outcome (*ex-post*) confidence in tax return preparation. An experimental methodology is used with 114 inexperienced and experienced participants that prepare an individual income tax return manually or with tax preparation software (a TDSS). Less experienced tax preparers have lower levels of *ex-ante* confidence, and are more likely to be overconfident in the accuracy of their performance. Furthermore, when examining only the participants that made errors in their tax return preparation task, those that prepare the return with the TDSS are significantly more likely to be overconfident in their performance. These results support the predictions of Noga & Arnold (2002) and suggest that inexperienced users' over-reliance on a TDSS (Masselli et al., 2002) may be due to individuals' overconfidence in the accuracy of their performance with the software.

INTRODUCTION

Tax return preparation software is considered to be an interactive, intelligent tax decision support system (TDSS) that assists tax preparers in their tax return preparation process (Masselli et al., 2002; Noga & Arnold, 2002). In recent years, the use of such software has skyrocketed.

Only 8% of returns were filed using personal tax preparation software in 1993, but by 2003, this number grew to 25% (Guyton et al., 2005; Toder, 2005). Even more dramatically, the Internal Revenue Service (IRS) recently reported that nearly 32 million taxpayers e-filed from their home computers in the 2009 filing season (IRS, 2009), nearly double the volume from 2006 (IRS, 2006). Similarly, tax professionals continue to rely on tax preparation software and e-filing to prepare their clients' returns (IRS, 2009).

Part of the popularity of the use of TDSS by novice and experienced tax preparers alike is that such software is perceived to increase tax preparers' accuracy in preparing their return, and thus may increase users' sense of *confidence* in their tax preparation abilities. A past advertising campaign by one software manufacturer, Intuit, included a direct appeal to users' confidence in its press release, claiming that its product (TurboTax) "Gives You Confidence Your Taxes Are Done Right" (Intuit, 2003). Such direct appeals to users' confidence levels have continued. During the beginning of the 2009 filing season, TurboTax's homepage included the claim that its software "lets you file with confidence", as well as guaranteeing "100% accuracy" for users (http://www.turbotax.com). Thus, tax preparation software such as TurboTax is marketed as giving users the confidence to accurately prepare a tax return. This suggests that confidence is an important part of the tax return preparation process, and that a TDSS assists with the process by increasing tax preparers' confidence levels.

Despite the potential importance of confidence in a multi-step, semi-structured, knowledge-based task such as tax return preparation, this factor remains an under-studied variable. While some psychology research has investigated the role of confidence in judgment and decision-making, such research does not always clearly distinguish between *ex-ante* confidence (prior to the actual performance of the task) and *ex-post* confidence (subsequent to

the actual task performance) (Bonner, 2008). Thus, it is unclear whether increases in *ex-ante* confidence would actually influence tax preparers' confidence in their ability to prepare an accurate tax return, and whether this same condition would be true with the introduction of a TDSS. Even more importantly, little is known regarding whether tax preparers' levels of *ex-post* confidence are appropriately calibrated. If the introduction of a TDSS indeed increases tax preparers' confidence levels, such increases in confidence could be detrimental if individuals develop a sense of *overconfidence* (i.e., greater perceived accuracy on a given task than the actual accuracy of the performance; Lichtenstein et al., 1982). The introduction of a TDSS may be particularly problematic because less experienced tax preparers often fail to function at the same level as more experienced preparers (Noga & Arnold, 2002).

The purpose of this study is two-fold. First, this study investigates the multi-faceted role of confidence in the tax return preparation process for tax preparers of varying experience levels, and for unaided (without a TDSS) and aided (using a TDSS) preparers. Second, this study deepens and extends conceptualizations of confidence by considering how this variable jointly functions as both an *input* and *output* of the decision process. This study investigates questions regarding the role of confidence in the tax return preparation process via an experimental methodology with 114 inexperienced and experienced tax preparers, each of whom prepare an income tax return either unaided (manually) or aided with a commercially available TDSS. Results indicate that levels of *ex-ante* confidence influence tax preparers' levels of accuracy in the aided (but not manual) condition. More experienced tax preparers had higher levels of *ex-ante* confidence, and were less likely to be overconfident in the accuracy of their performance. Interestingly, of the tax preparers that made errors in their tax return preparation, individuals using a TDSS were much more likely to be overconfident than manual preparers, supporting the

predictions of Noga & Arnold (2002). This phenomenon appears to be driven by the fact that TDSS users were more likely to make input errors that affected multiple facets of the tax return.

This paper contributes to the literature by considering how the introduction of a TDSS cognitively influences the tax return preparation process, and by extending existing theorizations of confidence to develop a more complete picture of this variable. This research is important for several reasons. First, most prior studies have examined either *ex-ante* or *ex-post* confidence, leaving open questions regarding the role of this variable in judgment and decision-making processes; this study suggests that confidence functions as both an input and output of decisions. Second, *overconfidence* is one of the most problematic biases in judgment and decision-making; results can be disastrous if individuals fail to realize the inaccuracy of their task performance (Plous, 1993). While some evidence suggests that decision aids may enhance overconfidence in *probabilistic* tasks (e.g., Davis & Kottemann 1994), this study indicates that the phenomenon of overconfidence continues in a complex, *non-probabilistic* task such as tax return preparation. Third, the explosion of the use of TDSS raises important questions regarding how such software influences the tax return preparation process, as well as whether such user perceptions of greater accuracy are beneficial or detrimental. The results of this study suggest that the finding in Masselli et al. (2002) that inexperienced tax preparers tend to over-rely on a TDSS is due to users' overconfidence in the *accuracy* of their performance when using tax preparation software. Thus, while the use of a TDSS does improve tax return preparation accuracy of both inexperienced and experienced tax preparers (Noga & Arnold, 2002), the results of this study indicate that *overconfidence* in performance with a TDSS may eventually lead to technology dominance for users that do not fully understand the task (Noga & Arnold, 2002; Arnold & Sutton, 1998).

The remainder of this paper is organized as follows. The second section presents prior literature and develops this study's hypotheses. The third section presents the research method for this experimental study. The fourth section provides an analysis of the results. The fifth and final section concludes with a discussion of the study's implications, limitations, and opportunities for future research.

PRIOR LITERATURE AND HYPOTHESES DEVELOPMENT

This study draws upon psychological theories of confidence in seeking to understand tax preparers' confidence in their tax preparation process. In general, confidence may be conceptualized in two different ways: as a *process* variable that is treated as a determinant of performance, or as an *outcome* variable that reflects individuals' confidence in their previous task performance (see Bonner, 2008). Process confidence refers to an individual's *ex-ante* (beforehand) confidence, which represents an individual's personal confidence in his or her ability to perform a task (Pincus, 1991; Whitecotton, 1996). Outcome confidence refers to an individual's *ex-post* (after the fact) confidence in performance (Einhorn & Hogarth, 1978; Lichtenstein et al., 1982). Exploring both of these roles is important in understanding the multi-faceted function of confidence. Prior literature on these conceptualizations of confidence is discussed below.

Process View of Confidence

According to Bonner (2008), the level of confidence that an individual brings to a task (*process confidence)* can have significant consequences on the quality of decision-making. However, Bonner (2008, p. 94) emphasizes that the effects of confidence on judgment and decision-making (JDM) are not well understood, since most studies "tend not to examine any effects of this confidence on subsequent JDM." Thus, while some prior studies in accounting

have examined confidence as a process variable in individuals' JDM, such studies still leave open questions regarding the influence of confidence on JDM.

The studies of Pincus (1991) and Whitecotton (1996) are examples of accounting research that have conceptually examined confidence as a process variable. As conceptualized by Pincus (1991), process confidence serves as a "stopping point" in auditor judgment, as auditors continue to gather necessary information until reaching a pre-established internal level of confidence. However, Pincus (1991) measured individual confidence in the decision *subsequent* to the actual decision, and did not specifically measure confidence as an input into the judgment process itself; thus, the results are unclear whether *ex-ante* confidence (prior to the task) would have influenced the task performance itself. Similarly, Whitecotton (1996) examined the relationship between an individual's personal confidence in the ability to perform a task (prior to engaging in the actual task) and reliance on a decision aid, finding a strong inverse relationship between personal confidence and decision aid reliance. Whitecotton (1996) also did not find a relationship between *ex-ante* confidence and performance accuracy. Thus, while Whitecotton (1996) established that personal confidence in ability level is an input in judgment tasks, results of the study suggest that process confidence may not directly influence performance accuracy for a given task.

Overall, few accounting studies have specifically examined the role of process confidence in the performance of a task. However, motivation theory suggests that higher levels of process confidence may lead to increased performance accuracy by increasing an individual's motivation to succeed in a given task (Benabou & Tirole, 2002). Similarly, lower levels of confidence may be detrimental to performance accuracy, because "when people expect to fail, they fail quite effectively" (Salancik, 1977; Benabou & Tirole, 2002, p. 873). While not

examining performance accuracy per say, sports psychology literature has specifically studied the degree to which "self-confidence" (i.e., process confidence) may influence competitive performance. Woodman & Hardy (2003, p. 443) define "self-confidence" as "one's belief in meeting the challenge of the task to be performed", and, in a meta-analysis of 48 studies on competitive performance, document a significant, large mean effect size for the positive relationship between self-confidence and performance.

The relationship between process confidence, motivation, and competitive performance suggests that individuals with higher levels of *ex-ante* confidence in their ability to perform a given task will have higher levels of task performance. It therefore follows that tax preparers with higher levels of *ex-ante* confidence in their ability to accurately prepare a tax return will perform this task with greater accuracy. This leads to the first hypothesis:

H_1: Tax preparers with higher levels of confidence in their ability to accurately prepare a tax return will demonstrate higher levels of accuracy in their performance of the task.

Individuals may differ in the levels of process confidence that they bring to a given task. One reason for such differences may be due to varying levels of *experience*. The theory of the *illusion of validity* postulates that greater experience with a task of moderate or high difficulty may increase an individual's confidence in his or her performance on that particular task due to the wider base of knowledge acquired by the individual (Einhorn & Hogarth, 1978). Over time, greater task experience logically leads to higher levels of process confidence. This suggests that tax preparers that have greater experience in preparing tax returns would be more likely to have greater *ex-ante* confidence in their ability to perform such a task. This leads to the second hypothesis:

H_2: *Tax preparers with greater tax return preparation experience will be more confident in their ability to accurately prepare a tax return than tax preparers with less tax return preparation experience.*

Outcome View of Confidence

In addition to a process variable, confidence may also be conceptualized as an *outcome variable*. Most studies in psychology that have examined confidence have used this conceptualization, assessing an individual's *ex-post* confidence in the performance of a particular task. Under this perspective, confidence is described as the degree of belief in the accuracy of task performance (Lichtenstein et al., 1982).

Outcome Confidence and the Role of Calibration

An evaluation of the appropriateness of an individual's level of *ex-post* confidence in the performance of a task can be made by comparing the level of confidence to the actual accuracy of a decision (Oskamp, 1965, 1982). Some psychologists label the correspondence between these constructs as the degree of an individual's *calibration*. Perfect calibration exists when an individual's self-assessed probability of the accuracy of his or her performance on a given task is precisely equal to the actual accuracy of the performance (Lichtenstein et al., 1982).

Perfect calibration is rare (Einhorn & Hogarth, 1978), resulting in overconfidence or underconfidence in performance. Overconfidence exists when an individual's self-assessed probability of the accuracy of performance on a task exceeds the actual accuracy, while underconfidence is characterized by greater accuracy in performance than self-assessed accuracy (Oskamp, 1965, 1982; Lichtenstein et al., 1982, 308).

Evidence of *overconfidence* is pervasive throughout the psychology literature (see Lichtenstein et al., 1982; McGraw et al., 2004). In general, while individuals may display

underconfidence on easy tasks, overconfidence is the most pronounced for tasks of moderate or high difficulty (Brenner et al., 1996). Overconfidence also abounds due to representativeness, ability, and internal coherence biases (Hogarth, 1980). Taken together, these prior psychology studies suggest that individuals are likely to be overconfident when assessing the accuracy of task performance of a task of moderate complexity, such as tax return preparation. The third hypothesis, a control hypothesis, is:

H_3: *Tax preparers will demonstrate overconfidence in their self-assessed accuracy of their preparation of a tax return.*

Understanding the factors that can influence individuals' levels of overconfidence is of critical importance, as, according to Plous (1993), "no problem in judgment and decision making is more prevalent and more potentially catastrophic than overconfidence" (Bonner, 2008, p. 93). Two such factors suggested by prior research that may influence the level of overconfidence are *experience* and the use of a *decision aid*, such as a TDSS.

Outcome Confidence and Experience

Several psychological studies have explored the relationship between experience and overconfidence. Some prior literature has suggested that individuals with greater task experience are more prone to overconfidence. Over time, greater task experience may lead to increased levels of *ex-post* confidence in performance that could exceed the corresponding improvements in performance accuracy (Einhorn & Hogarth, 1978; Griffin & Varey, 1996). For instance, Arkes et al. (1986) found that individuals with a higher knowledge base in a particular domain were more prone to overconfidence in their performance than those with a moderate level of knowledge, as measured by the degree of miscalibration between the self-assessed accuracy of performance and actual performance.[1]

Alternatively, other studies have found that experienced individuals may be less prone to overconfidence than those with little task experience. Oskamp (1962) found that less-experienced decision makers had greater degrees of miscalibration than inexperienced performers. Individuals with greater experience may have higher self-awareness of the potential accuracy of their performance (Kahneman & Tversky, 1982). Others suggest that the degree of calibration, or the relationship between confidence levels and accuracy, is strongest for experienced performers performing relatively easy tasks who receive complete feedback regarding the accuracy of their decisions (Fischer & Budescu, 2005). Thus, experts in a particular domain may be less prone to overconfidence than novices due to greater task experience (Keren, 1987).

Closer analysis shows a common theme on the relationship between task experience and overconfidence. The theory of the illusion of validity states that overconfidence is the most pronounced for tasks of moderate or high difficulty, while individuals performing easier tasks generally suffer from underconfidence (Einhorn & Hogarth, 1978). As an individual's performance on a given task generally determines whether the task is difficult or easy, it follows that an individual that performs well on a particular task is less likely to be overconfident due to the task's ease (Brenner et al., 1996; Fisher & Budescu, 2005). Individuals with greater task experience are likely to demonstrate improved calibration as compared to individuals with less task experience due to the reduced difficulty level of the task. Likewise, tax preparers with greater task experience are likely to demonstrate improved calibration as compared to less-experienced tax preparers. This leads to the fourth hypothesis:

H_4: Tax preparers with less tax return preparation experience will demonstrate higher levels of overconfidence in their self-assessed

accuracy of their tax return preparation than tax preparers with more tax return preparation experience.

Outcome Confidence and Decision Aids

An additional variable that may be associated with overconfidence is the use of a decision aid. Evidence is mixed regarding whether participants are more or less likely to exhibit overconfidence when using a decision aid to perform a task. Indeed, Rose (2000) calls for future studies to investigate alternative theories and settings to understand the conditions in which overconfidence when using a decision aid may be present.

On the one hand, the interaction between individuals and computerized machines, including decision aids, has been characterized as simulating group interaction (Woods & Roth, 1988; Kasper, 1996). Others have suggested that an intelligent decision aid approximates an "electronic colleague", and that collaboration between such an aid and the user is similar to the type of interaction between a two-person group (Arnold & Sutton, 1998). Thus, the interaction of multi-person groups is a useful analogy in describing the interaction between individuals and collaborative decision aids.[2] In general, the interactive nature of groups leads to group decisions that are typically even more confident and accurate in judgments than individual decisions (Sniezek & Henry, 1989). Other studies have demonstrated that groups tend to be better calibrated in the confidence of their decisions than individual decision makers (e.g., Ahlawat 1999). Thus, some evidence suggests that groups may be *less* prone to overconfidence than individuals, implying that overconfidence may be less prevalent among individuals using decision aids to perform a task.

On the other hand, several studies in the systems literature have documented that the use of such aids may *increase* users' overconfidence in the accuracy or quality of such decisions

(e.g., Davis & Kottemann, 1994; Kottemann et al., 1994; Kahai et al., 1998). This overconfidence may be due to the illusion of control (Langer, 1975) that individuals exhibit when overvaluing the usefulness of a decision aid to perform a probabilistic task. Furthermore, as suggested by Kasper (1996), interaction with a decision aid may increase user overconfidence due to the publicized expertise of such a system, particularly in regards to the system's inquirability (cuing the user to particular alternatives or decisions). Thus, compared to unaided users, decision aid users may overweight the extent to which their performance increases with such an aid, and may be more likely to be overconfident in their task performance.

Overall, using a decision aid tends to improve the accuracy of performance, but also increases users' *ex-post* confidence in the accuracy of their performance. Whether the higher degree of confidence from using a decision aid such as a TDSS would exceed the increased accuracy resulting from the use of such an aid is unclear. Specifically, tax preparers may demonstrate increased accuracy in performance when using a TDSS (Noga & Arnold, 2002), but the increased confidence brought about by the use of the TDSS may outweigh improvements in performance. Thus, the following two-sided hypothesis is proposed:

H_5: *Tax preparers will have differing levels of overconfidence in the accuracy of their tax return preparation based on whether they prepare the return manually or with tax preparation software.*

RESEARCH METHOD

To examine the five hypotheses, an experiment was conducted using a between-subjects design with two treatment conditions. The section below details the experimental task, participants, experimental procedures, and the operational measures of variables.

Experimental Task

The experimental task consisted of the completion of an individual tax return for a hypothetical family either with a commercially available TDSS (TurboTax, a type of tax preparation software) or manually with paper-based forms. Tax preparation software represents a TDSS with high external validity due to its widespread commercial availability (Masselli et al., 2002; Noga & Arnold, 2002). This experimental task itself was based on the complex return developed in Noga & Arnold (2002), and was modified to include a married couple filing jointly with two dependents, W-2 wages, itemized deductions, a child tax credit, Schedule C income, and self-employment tax.

All participants were randomly assigned to two treatment groups: the aided group (i.e., with the TDSS) and the unaided group (i.e., the manual group). Participants in both the aided and unaided condition received the same basic taxpayer information and supporting documentation. Following Noga & Arnold (2002), participants in the unaided condition also received the following forms and instructions: 1040, 1040-A, 1040-EZ, Schedule A, Schedule B, Schedule C, Schedule C-EZ, Schedule D, Self-Employment Tax, and Earned Income Credit. As with Noga & Arnold (2002), additional forms were provided to this group to determine whether participants were capable of selecting the appropriate forms. The instructions for this task were adapted from Noga & Arnold (2002) and Masselli et al. (2002).

Participants

The participants in this research study were students from a large southeastern university. Two groups of students participated in the study: students currently enrolled in the first undergraduate course in taxation, and students currently enrolled in a graduate-level taxation course. A total of 69 students were enrolled in the graduate taxation course, and 114 students were enrolled in the undergraduate taxation course. Originally, 132 participants volunteered to

participate in the study; the responses of 114 participants could be used in the final analysis. Of the 18 participants that could not be included, five volunteers did not report to the experiment, six participants did not complete all required experimental materials, three participants either failed the manipulation check or did not answer the question, three responses had corrupted electronic files that could not be read at a later date, and one participant completed the experimental materials out of order. Of these 114 participants, 39 were students currently enrolled in the graduate-level taxation course, while 75 of the participants were currently enrolled in the undergraduate-level course. Among participants from the graduate level course, 41 percent were master's of accounting students, 33.3 percent were upper-level undergraduate accounting students, 23 percent were master's of taxation students, and 2.6percent were MBA students. By comparison, 80 percent of the participants from the undergraduate course were accounting majors, 13 percent were finance majors, and the remainder were other business majors.

Table 1 summarizes information regarding the two groups of participants. The more experienced group (students in the graduate-level taxation course) was significantly (at $p < 0.05$) older, had a higher grade point average (GPA), and had completed more taxation courses. Participants in the more experienced group were also more likely to prepare both their own tax return and returns for third parties, and had completed a higher volume of tax returns in the past than the inexperienced group. While the more experienced group reported greater familiarity with TurboTax software, groups did not differ (at $p < 0.05$ significance) in whether they had used tax software in the past, or in their familiarity with the current TurboTax advertising campaign. Groups also did not display any significant difference in prior tax work experience. As a result, these groups serve as proxies for tax preparers with varying levels of experience.[3]

[Insert Table 1 about here]

Experimental Procedures

Each participant in the inexperienced or experienced group was randomly assigned to one of the treatment conditions (aided or unaided) and was informed of the location of their assigned session. To maximize participation, five different experimental sessions were held. Each session consisted of both an aided and unaided group, held in separate locations. Refer to Table 2 for a breakdown of the participants.[4]

[Insert Table 2 about here]

Participants in the unaided condition completed the experiment in a traditional classroom environment. The researcher or an assistant provided a general overview of the study and distributed a set of four packets to each participant that were to be completed in a pre-specified order. Each packet contained an experimental task and a legal-sized envelope. Participants were advised to open each packet, complete the experimental task, and seal the completed task inside the legal-sized envelope before proceeding to the next packet. At the end, the participants turned in the four, sealed, legal-sized envelopes and all of the other experimental materials.

The first packet contained a pre-experimental questionnaire to obtain (1) information regarding participants' background experience in tax preparation and (2) an assessment of participants' confidence in their ability to prepare a tax return either with or without tax preparation software, as well as their confidence in the accuracy of the tax preparation software itself. The second packet contained the experimental task itself, and included taxpayer information, supporting documentation, forms, and instructions. This task was designed to measure participants' accuracy in their tax preparation decisions. The third packet contained a post-experimental questionnaire gathering additional demographic information, participants'

confidence in their performance of the task, and their confidence in the accuracy of tax preparation software. The fourth packet contained a copy of the correctly prepared tax return (determined via the consensus of the researcher and two taxation professors). This final packet also contained a short questionnaire with questions measuring participants' future confidence in their ability to complete a task both with and without tax preparation software, and the manipulation check.

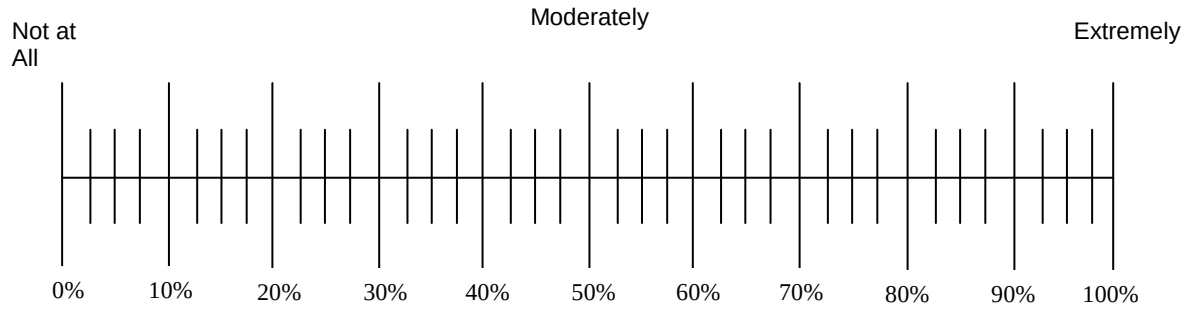
Participants in the aided condition completed the experiment in a computer-laboratory setting. Their packets were identical to those in the unaided group, with the exception of the material in the second packet. The participants in the aided group did not have hard copies of the forms or instructions, but instead received a floppy diskette and copies of brief instructions of how to launch TurboTax in the laboratory. Furthermore, these participants completed the experimental task using TurboTax, and then sealed the floppy diskette with a saved copy of the prepared return in one of the legal-sized envelopes.

The experiment was pre-tested by nine graduate students. Based on feedback received by pre-testers, some changes in wording were made to questionnaire items.

Measurement of the Variables

One dependent variable of interest was *performance accuracy* (H_1). This was measured as: (1) the number of errors made in the completion of the experimental task, and (2) the absolute value of the dollar amount of the errors (Noga & Arnold, 2002).

Confidence is the multi-dimensional construct of interest. Both aspects of "confidence" were measured on a 100 point-scale (Whitecotton, 1993; Whitecotton, 1996) ranging from "not at all" (0%), "moderately" (50%), and "extremely" (100%). Participants used the following scale to answer the confidence questions:



Measures of *ex-ante* confidence in performance (H_1 and H_2) were obtained prior to the experimental task itself and were measured based on responses to the following questions: "How confident are you in *your ability* to accurately prepare an individual income tax return manually *without* the assistance of tax software?" and "How confident are you in *your ability* to accurately prepare an individual income tax return *with* the assistance of tax software?"

The measurement of outcome confidence, or *ex-post* overconfidence (H_3, H_4, H_5) concerns the *miscalibration of errors*, which is operationalized as the *difference between the number of errors* participants believed they made in the preparation of the tax return and the actual number of errors (see Brenner et al., 1996). Participants may be over- or under-confident in this measure. However, the scale used to measure this item gave participants the option to select "five or more errors", which biases the measure toward calibration since participants who made extensive errors could select this option and still be classified as accurately calibrated. Concerns about this bias, however, may be partially remedied by the fact that the measures of participants' *ex-post* confidence in their accuracy of performance and the self-assessed number of errors are very highly correlated ($r = -.779$, $p < 0.001$).

Finally, taxation course enrollment serves as a proxy for *tax return preparation experience* (H_2, H_4). Students currently enrolled in the undergraduate taxation course are considered inexperienced tax preparers, while students currently enrolled in the graduate taxation

course are considered experienced tax preparers. Supplemental analyses use the number of tax returns previously prepared as an alternative measure of tax return preparation experience.

RESULTS

Data collected from the experimental procedures are used to test the five hypotheses pertaining to process confidence and outcome confidence in tax return preparation. Descriptive statistics are presented in Table 3, Panels A through D. Specifically, this table presents means and standard deviations for measures of *ex-ante* confidence, performance accuracy as measured by the number of errors and absolute value of the dollar amount of errors, and miscalibration in confidence (i.e., underconfidence or overconfidence). The table details the breakdown of these variables at both the group level (experienced or inexperienced) and the treatment condition (aided or unaided). Examination of these descriptive statistics demonstrates that the experienced group had higher levels of *ex-ante* confidence in their ability to prepare a tax return manually, had higher levels of performance accuracy in the task itself, and were less likely to be overconfident in the accuracy of their tax return preparation than the inexperienced group. Both inexperienced and experienced groups had higher levels of *ex-ante* confidence in their tax return preparation abilities under aided rather than unaided conditions. Furthermore, for the inexperienced taxpayers, the aided group had higher levels of performance accuracy in obtaining an accurate tax liability than the unaided group.

[Insert Table 3 about here]

Process View of Confidence – H_1 and H_2

Hypothesis 1 predicts that tax preparers' *ex-ante* confidence in their ability to accurately prepare a tax return (both with and without a TDSS) will be positively associated with performance accuracy (as measured by both the number and value of errors). Thus, four

regressions are run to separately assess the aided and unaided groups and the two measures of performance accuracy. Control variables for the number of tax returns previously prepared, tax work experience, and taxation classes are also included. As shown in Table 4, results differ between the unaided and aided groups. For the unaided (manual) group, the relationships between *ex-ante* confidence in performance *without* software and both the number and dollar value of errors are insignificant. Thus, increased confidence does not improve the manual preparation of tax returns. However, the control variable for the number of tax returns previously prepared is statistically significant ($p < 0.01$) in the analysis of the number of errors made by the unaided group; participants who have prepared more returns in the past make fewer errors.

[Insert Table 4 about here]

In the analysis of the aided condition, an additional control variable is included for participants' familiarity with TurboTax (the TDSS in the study); *ex-ante* confidence in performance *with* software is the independent variable in question. For participants completing a tax return with a TDSS, higher levels of *ex-ante* confidence in performance result in fewer errors in both magnitude and dollar value (i.e., higher levels of performance accuracy) at a statistically significant level ($p < 0.05$). Participants' levels of confidence appear to be a particularly important influence on performance accuracy when using a TDSS. None of the control variables are statistically significant. Thus, Hypothesis 1 is supported for participants in the aided condition.

Hypothesis 2 examines whether tax preparers with greater tax return preparation experience will be more confident in their ability to accurately prepare a tax return. To test this hypothesis, two separate ANOVAs are used with the dependent variables of *ex-ante* confidence in performance without a TDSS (manually), and *ex-ante* confidence in performance with a

TDSS; the independent variable for both analyses is the inexperienced or experienced group.[5] As shown in Table 5, Panel A, tax preparers in the more experienced group were significantly more confident in their ability to prepare a tax return manually ($p < 0.001$). Interestingly, there were no such differences between the groups in their levels of *ex-ante* confidence in their ability to prepare an accurate tax return with software (TDSS) as shown in Table 5, Panel B. This seems to suggest that inexperienced tax preparers believe that they can use a TDSS to perform at a higher level of accuracy, equivalent to a more experienced tax preparer.

[Insert Table 5 about here]

An alternative reason for the difference in the effect of experience on confidence levels is that both groups demonstrated low levels of familiarity with tax preparation software. A supplemental analysis uses the number of tax returns previously prepared as an alternative measure of tax return preparation experience, while controlling for tax work experience and taxation classes; the analysis of *ex-ante* confidence when using a TDSS also controls for familiarity with TurboTax. Table 5, Panel C shows the regression analyses that demonstrate that prior tax return preparation experience (i.e., the number of returns previously prepared) significantly influences *ex-ante* confidence both with and without the TDSS (both $p < 0.01$, one-tailed). Greater familiarity with the software also increases confidence in the ability to prepare a return with a TDSS ($p < 0.05$). Overall, these results generally support H_2 by indicating that tax preparers with greater tax return preparation experience are more likely to have greater degrees of *ex-ante* confidence in tax return preparation; *direct* task experience is particularly important.

Outcome View of Confidence – H_3, H_4, and H_5

Hypothesis 3 is a control hypothesis used to establish that tax preparers will demonstrate overconfidence in their self-assessed accuracy of their preparation of a tax return. An

examination of the descriptive statistics in Table 3, Panel D demonstrates that on average, participants are overconfident in their assessment of their performance, in that the actual number of errors in preparation exceeds their self-estimated number of errors. A t-test demonstrates that the overall measure of overconfidence is significantly different from zero ($p < 0.001$). Thus, Hypothesis 3 is supported, in that participants are overconfident in their self-assessed tax return preparation accuracy.

Hypothesis 4 predicts that tax preparers with greater tax return preparation experience will demonstrate higher levels of overconfidence in their self-assessed performance accuracy, while Hypothesis 5 predicts that tax preparers' level of overconfidence in tax return preparation will vary based on whether they prepare the return manually or with software. Thus, these hypotheses examine the differences in participants' estimated number of errors and the actual number of errors (i.e., overconfidence) to determine if any of the conditions were more prone to overconfidence. These hypotheses are tested with an ANOVA. Results shown in Table 6, Panel A indicate that participants in the inexperienced group are significantly more overconfident than those in the experienced group ($p < 0.01$), supporting the predictions of H_4.[6] However, no significant differences emerge between participants in the aided and unaided conditions.

[Insert Table 6 about here]

Further analysis is performed to determine whether participants that completed the task accurately (i.e., without any errors) may be confounding the results for H_5. Using a TDSS does appear to improve tax preparer accuracy, as 79 percent of tax preparers made errors on the return when preparing it manually, compared to only 41 percent of tax preparers that used a TDSS (significantly different at $p < 0.001$). This suggests that overconfidence may be particularly

problematic for tax preparers that were unable to perform the task accurately. Thus, additional analysis eliminates the participants that completed the task without any errors.

Eliminating the participants that completed the task accurately results in 46 participants in the unaided condition and 23 participants in the aided condition that made errors in their preparation. A supplemental ANOVA considers whether there are any differences in overconfidence between aided and unaided users that made errors in their performance. Considering only the participants that made errors, those that completed the tax return using a TDSS have significantly higher levels of overconfidence than those that completed the tax return manually ($p < 0.001$) as shown in Table 6, Panel B. Thus, using a TDSS appears to result in a false sense of overconfidence among users who make errors in their tax return preparation. Hypothesis 5 is therefore partially supported.

Additional Analysis

Additional analysis further explores the consequences of tax preparers' overconfidence. One such area pertains to the *type* of errors made by participants in the aided and unaided conditions. For participants that made errors in their tax return preparation, no statistically significant differences between aided and unaided participants emerged regarding the net magnitude of such errors. However, closer investigation reveals that certain types of errors were more prevalent for paper-based returns than ones completed with software. Table 7 provides a summary of the differences in errors made by participants in the aided and unaided conditions. Overall, tax preparers using a TDSS were significantly more likely to *omit* certain items from the tax return at a statistically significant level (all $p < 0.05$, two-tailed), including information about dependents, wages, and itemized deduction items on Schedule A. Such omissions for participants in the TDSS condition were likely a result of answering a question incorrectly in the electronic

"interview" in the TDSS, which likely affects multiple items on the return. For example, a taxpayer who omits information regarding a dependent will pay more tax, due to the impact of this error on both dependency exemptions and the child tax credit. Thus, errors when using tax preparation software may be magnified across the return.

[Insert Table 7 about here]

Other types of errors were more prevalent between the aided and unaided conditions. Tax preparers using the TDSS were more likely to calculate an incorrect amount for business expenses on the Schedule C-EZ, in large part due to the common tendency to treat mortgage interest (a personal expense) as a business expense. This tendency may have been more likely due to the electronic prompting of the TDSS interview, causing participants to misclassify the item to a much larger degree than in the manual condition. On the other hand, participants in the manual condition were much more likely to make errors in the calculation of tax amounts. In fact, of the participants making an error in the manual condition, 76 percent either neglected to compute or made an error in the calculation of the child tax credit, resulting in an overpayment of tax. Overall, while participants were much less likely to make errors in their tax return preparation when using the TDSS, those that *did* make errors tended to have mistakes whose consequences were magnified across the return, or were a consequence of the nature of the interactive questionnaire in the TDSS. This helps to explain why tax preparers may develop a false sense of confidence when using a TDSS.

Another area of further investigation concerns the factors associated with overconfidence. While this study focuses on the associations between inexperience and a TDSS and overconfidence, other variables may also be important in explaining this tendency. One such important factor is gender, as some evidence exists that men may be more prone to

overconfidence, particularly in traditionally "masculine" fields (Bonner, 2008). Supplemental analysis (not tabulated) shows that this gender difference is also prevalent in the realm of tax return preparation, with male participants significantly more likely to be overconfident than female participants ($p < 0.05$, two-tailed). Further investigation shows that of the participants who made errors in their tax return preparation, male participants using the TDSS were particularly prone to overconfidence ($p = 0.06$). Thus, addressing the potential effects of overconfidence is particularly important for male tax preparers.

DISCUSSION AND CONCLUSION

This study investigates the role of confidence in tax return preparation, and examines tax preparers' process (*ex-ante*) and outcome (*ex-post*) confidence in preparing a tax return both manually (without a TDSS) and with tax preparation software (with a TDSS). Hypothesis 1 investigates whether tax preparers with higher levels of *ex-ante* confidence have higher levels of accuracy in preparing a tax return. Results demonstrate that this relationship is present for tax preparers using a TDSS, but not for those performing the task manually. Hypothesis 2 examines whether greater experience increases levels of *ex-ante* confidence; results using prior tax return preparation experience support this prediction. Thus, this study adds to the literature by examining how confidence may affect subsequent judgments and decisions (Bonner, 2008), by demonstrating that *ex-ante* confidence is particularly important when preparing a tax return with software.

Hypothesis 3 tests whether tax preparers are subject to miscalibration in the form of overconfidence; results support this prediction, indicating that tax return preparation is another field where the "potentially catastrophic" problem of overconfidence arises (Plous, 1993). Whereas prior studies have tended to focus on domains whose outcome was at least partially

probabilistic in nature (e.g., Davis & Kottemann, 1994), these results show that overconfidence is a problem even in non-probabilistic domains such as tax return preparation. Hypothesis 4 examines whether experience is related to levels of overconfidence. Results indicate that overconfidence abounds among individuals with *less* prior tax return preparation experience, supporting the view that when a task is of greater difficulty for an individual, overconfidence is more likely (Einhorn & Hogarth, 1978).

Finally, Hypothesis 5 addresses whether overconfidence is more likely in tax return preparation when using a TDSS. While there were no differences in overconfidence levels for the whole sample between the aided and unaided conditions, an interesting result is that of the participants that made errors in their tax return preparation, participants using software were significantly more likely to be overconfident in the perceived accuracy of their return. Thus, while confidence and accuracy both tend to increase in an interactive, computerized environment (Sniezek & Henry, 1989), tax preparers that are inaccurate in preparing returns using software are more likely to be overconfident in performance and not perceive the full extent of their errors. This also confirms the speculation of Noga & Arnold (2002) that tax preparation software may give the tax preparer a false sense of confidence in the accuracy of the results, which may make them even less likely to be aware of any potential errors.

These findings have both theoretical and practical implications. The pervasive nature of overconfidence (Lichtenstein et al., 1982; McGraw et al., 2004) is shown to exist in a complex, multi-step task such as task return preparation. That overconfidence in inaccurate performance is more likely when preparing a tax return using software suggests that the "illusion of control" phenomenon present when using a DSS (e.g., Davis & Kottemann 1994) may also be present in the performance of a non-probabilistic based task. Finally, the results of this study suggest that

the over-reliance on tax preparation software demonstrated by novice taxpayers in Masselli et al. (2002) may be due to individuals' overconfidence in their accuracy of performance with the software. Over time, this overconfidence may lead to technology dominance for users that do not fully understand the task (Arnold & Sutton, 1998), and explains why individuals that are not accustomed to performing a task in a manual environment may be unable to do so (Noga & Arnold, 2002).

Several practical implications also arise, particularly in highlighting the differences in tax return preparation in a manual or computerized DSS environment. That overconfidence in inaccurate performance is more likely when a DSS is used suggests that tax preparers may not fully understand the impact of errors. For instance, in the case of tax preparation software, incorrectly answering an item in the software's questionnaire (such as information on dependents) may affect multiple items on the resulting tax return (such as dependency exemptions and the child tax credit, resulting in the tax preparer overstating the tax liability). Thus, errors in preparation may be magnified when tax preparation software is used. This deficiency suggests an opportunity for educators to provide extensive training on the use of tax preparation software, particularly on the impact of errors on the final return. Furthermore, because increased confidence in performance with software was linked with more accurate performance, training on the use of such software in order to properly calibrate users' confidence levels is critical. Overall, tax preparers should not abandon the use of TDSS, but should be adequately trained regarding the systems' limitations and the potential for pervasive errors in order to reduce overconfidence.

The implications of the current study should be interpreted in light of its limitations. First, this study used student subjects as a proxy for inexperienced and experienced tax preparers.

However, these groups still serve as proxies for tax preparers with varying levels of tax return preparation experience; furthermore, when applicable, alternative proxies for "experience" were also employed. Second, the scale that measured participants' estimated number of errors made in the experimental task contained an option for "five or more errors," which biases the measure toward calibration (i.e., against finding overconfidence). Despite this limitation, all experimental conditions still displayed a tendency toward overconfidence, and participants' estimates of their predicted number of errors were significantly correlated with *ex-post* confidence in performance. Third, some variables are operationalized with single self-reported measurements.

The limitations and implications of the current study suggest several avenues for future research. Future studies could determine the effects of training in the use of tax preparation software on tax preparers' confidence levels. Another possibility is to determine if offering participants incentives for appropriate calibration could reduce overconfidence. Finally, researchers could examine the constructs of trust or face validity in relation to the TDSS to determine if such facets are necessary precursors of confidence. Such future research can begin to examine some of the issues raised in this study.

ENDNOTES

[1] The findings of this particular study, however, should be interpreted with caution since the group with higher domain knowledge also was much less likely to rely upon a decision aid with high predictive accuracy. It is possible that the greater knowledge base of the expert group was not the sole determinant of the group's overconfidence (Arkes et al. 1986).

[2] While the interaction between an individual user and a computerized decision aid is a type of group processing (Woods & Roth, 1988), judgments regarding confidence in the performance of a task will reflect the judgment of only one member of the group: the individual user. Nevertheless, the analogy of the group is a useful tool in understanding this phenomenon.

[3] In order to motivate participants to attend to the experimental task, students in both the undergraduate and graduate taxation classes were awarded 10 extra-credit points (approximately 2% of the final course grade) for making a good faith effort to complete the study without obvious or blatant carelessness. Nearly all participants appeared to be highly motivated and focused on the task, and instructors in each of the classes reported that the participants were interested in the outcome.

[4] There were no statistically significant differences in any demographic or tax experience questions between subjects assigned to the aided or unaided conditions (all $p > 0.10$).

[5] Separate ANOVAs, rather than a MANOVA, are used because the two dependent variables are very highly correlated (correlation coefficient = 0.601).

[6] Results are robust to an alternative specification of "experience," as participants who had previously prepared more (fewer) tax returns in the past were also less (more) prone to overconfidence ($p < 0.01$).

ACKNOWLEDGEMENTS

I am very grateful for the helpful comments and suggestions from Vicky Arnold, Donna Bobek Schmitt, Charles Cho, Clark Hampton, Charlie Kelliher, Steve Sutton, and participants at the 2007 AAA Information Section JIS Research Development Workshop. I am also grateful for the assistance in carrying out the experimental procedures provided by Joe Canada and Will Hageman.

REFERENCES

Ahlawat, S. S. (1999). Order effects and memory for evidence in individual versus group decision making in auditing. *Journal of Behavioral Decision Making, 12,* 71-88.

Arkes, H. R., Dawes, R. M., & Christensen, C. (1986). Factors influencing the use of a decision rule in a probabilistic task. *Organizational Behavior and Human Decision Processes, 37,* 93-110.

Arnold, V. & Sutton, S. (1998). The theory of technology dominance. *Advances in Accounting Behavioral Research, 1,* 175-194.

Benabou, R. & Tirole, J. (2002). Self-confidence and personal motivation. *Quarterly Journal of Economics, 117*(3), 871-915.

Bonner, S. (2008). *Judgment and decision making in accounting*. Upper Saddle River, NJ: Pearson Prentice Hall.

Brenner, L. A., Koehler, D. J, Liberman, V. & Tversky, A. (1996). Overconfidence in probability and frequency judgments: A critical examination. *Organizational Behavior and Human Decision Processes, 65*(3), 212-219.

Davis, F. D. & Kottemann, J. E. (1994). User perceptions of decision support effectiveness: Two production planning experiments. *Decision Sciences, 25*(1), 57-78.

Einhorn, H. J. & Hogarth, R. M. (1978). Confidence in judgment: Persistence of the illusion of validity. *Psychological Review, 85*(5), 395-416.

Fischer, I. & Budescu, D. V. (2005). When do those who know more also know more about how much they know? The development of confidence and performance in categorical decision tasks. *Organizational Behavior and Human Decision Processes, 98,* 39-53.

Griffin, D. W. & Varey, C. A. (1996). Towards a consensus on overconfidence. *Organizational Behavior and Human Decision Processes, 65*(3), 227-231.

Guyton, J. L., Korobow, A. K., Lee, P.S., & Toder, E. J. (2005). The effects of tax software and paid preparers on compliance costs. *National Tax Journal, 58*(3), 436-448.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis*. Upper Saddle River, NJ: Pearson Education, Inc.

Hogarth, R. (1980). *Judgment and choice*. New York, NY: John, Wiley & Sons, Ltd.

Internal Revenue Service (IRS). (2006). On-line filing sets new record; More than 17 million returns e-filed by home computer users. *IR-2006-64* (April 14). Available at http://www.irs.gov/newsroom/article/0,,id=156222,00.html

Internal Revenue Service (IRS). (2009). Two Out of Three Individuals Now Using IRS E-File. IR-2009-99 (November 2). Available at http://www.irs.gov/newsroom/article/0,,id=214853,00.html

Intuit. (2003). TurboTax for 2003 Gives You Confidence Your Taxes Are Done Right. Press Release (December 2). Available at http://web.intuit.com/about_intuit/press_releases/2003/12-02a.html

Kahai, S. S., Solieri, S. A., & Felo, A. J. (1998). Active involvement, familiarity, framing, and the illusion of control during decision support system use. *Decision Support Systems, 23*(2), 133-148.

Kahneman, D. & Tversky, A. (1982). Intuitive prediction: Biases and corrective procedures. In: D. Kahneman, P. Slovic, & A. Tversky (Eds), *Judgments under Uncertainty: Heuristics and Biases* (pp. 414-421). New York: Cambridge University Press.

Kasper, G. M. (1996). A theory of decision support system design for user calibration. *Information Systems Research, 7*(2), 215-232.

Keren, G. (1987). Facing uncertainty in the game of bridge: A calibration study. *Organizational Behavior and Human Decision Processes, 39*(1), 98-114.

Kottemann, J. E., Davis, F. D. & Remus, W. E. (1994). Computer-assisted decision making: Performance, beliefs, and the illusion of control. *Organizational Behavior and Human Decision Processes, 57*(1), 26-37.

Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology, 32*, 311-328.

Lichtenstein, S., Fischhoff, B. & Phillips, L. D. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, &. A. Tversky. (Eds), *Judgment under Uncertainty: Heuristics and Biases* (pp. 306-334). Cambridge: Cambridge University Press.

Masselli, J. J., Ricketts, R. C., Arnold, V. & Sutton, S. G. (2002). The impact of embedded intelligent agents on tax-reporting decisions. *Journal of the American Taxation Association, 24* (2), 60-78.

McGraw, P. A., Mellers, B. A. & Ritov, I. (2004). The affective costs of overconfidence. *Journal of Behavioral Decision Making, 17*, 281-295.

Noga, T. & Arnold, V. (2002). Do tax decision support systems affect the accuracy of tax compliance decisions? *International Journal of Accounting Information Systems, 3,* 125-144.

Oskamp, S. (1965). Overconfidence in case-study judgments. *The Journal of Consulting Psychology, 29,* 261-265.

Oskamp, S. (1982). Overconfidence in case study judgment. In: D. Kahneman, P. Slovic, & A. Tversky (Eds), *Judgments under Uncertainty: Heuristics and Biases* (pp. 287-289). New York: Cambridge University Press.

Pincus, K. V. (1991). Audit judgment confidence. *Behavioral Research in Accounting, 3*, 39-65.

Plous, S. (1993). *The psychology of judgment and decision-making*. Philadelphia: Temple University Press.

Rose, J. (2000). Behavioral decision aid research: Decision aid use and effects. In: V. Arnold & S. Sutton (Eds), *Researching Accounting as an Information Systems Discipline* (pp. 111-133). Sarasota, FL: American Accounting Association.

Salancick, G. (1977). Commitment and the control of organizational behavior and belief. In: B. Staw & G. Salancik (Eds), *New Directions in Organizational Behavior* (pp. 1-54). Chicago, St. Clair Press.

Sniezek, J. A. & Henry, R. A. (1989). Accuracy and confidence in group judgment. *Organizational Behavior and Human Decision Processes, 43*(1), 1-28.

Toder, E. J. (2005). Changes in tax preparation methods, 1993-2003. *Tax Notes, 100*(6), 759.

Whitecotton, S. M. (1993). *Decision aid reliance as a determinant of accuracy in earnings forecasting judgments: The impact of contextual and decision maker characteristics*. Unpublished Ph.D. Dissertation, University of Oklahoma.

Whitecotton, S. M. (1996). The effects of experience and confidence on decision aid reliance: A causal model. *Behavioral Research in Accounting, 8,* 194-216.

Woodman, T. & Hardy, L. (2003). The relative impact of cognitive anxiety and self-confidence upon sport performance: A meta-analysis. *Journal of Sports Sciences, 21*, 443-457.

Woods, D. D. & Roth, E. (1988). Cognitive engineering: Human problem solving with tools. *Human Factors*, *30*(4), 415-430.

Table 1
Demographic Information

	Inexperienced (n = 75)	**Experienced (n = 39)**
Average Age	**24.23 Years***	**27.50 Years***
Gender (Females)	56%	67%
Average Self-Reported GPA	**3.33***	**3.48***
Average Tax Classes Taken	**1.19***	**2.44***
Prepare Tax Return – Own	**45%***	**82%***
Prepare Tax Return – Other Parties	**31%***	**51%***
Average Tax Work Experience	2.53 Months	5.56 Months
Average Tax Returns Prepared[a]	**1.29***	**2.18***
Previously Used Tax Software	49%	67%
Familiarity with TurboTax[b]	**1.95***	**2.74***
Familiarity with TurboTax Advertising[b]	2.29	2.21

[a]6-point scale, measured with 0 = None; 1 = 1-5 returns; 2 = 6-10 returns; 3 = 11-15 returns; 4 = 16-20 returns; and 5 = More than 20 returns.
[b]7-point scale, measured with 1 = Not at All and 7 = Very.

*Significantly different between the two groups at $p < 0.05$ (two-tailed)

Table 2
Number of Participants

	Unaided (Manual) Group	Aided (TDSS) Group	Overall
Inexperienced	39	36	75
Experienced	19	20	39
Overall	58	56	114

Table 3
Descriptive Statistics

Group	Unaided Mean (Std. Dev)	Aided Mean (Std. Dev)	Overall Mean (Std. Dev)
Panel A: Process Confidence (*Ex-Ante*)—Prior to using tax preparation software[a]			
Inexperienced	**52.99**[*],[**] (26.30)	**76.65**[**] (22.05)	64.82 (21.84)
Experienced	**70.13**[*],[**] (18.23)	**81.46**[**] (20.26)	75.80 (16.86)
Overall	58.85 (25.13)	78.30 (21.49)	68.57 (20.87)
Panel B: Performance Accuracy of Participants (Absolute Value of Dollar Amount of Errors)			
Inexperienced	**6258.56**[*],[**] (11592.22)	**1260.00**[**] (2415.33)	3999.65 (8876.19)
Experienced	**924.37**[*] (893.11)	596.95 (1078.20)	756.46 (1078.20)
Overall	4692.71 (9842.65)	1023.20 (2082.87)	2890.14 (7373.88)
Panel C: Performance Accuracy of Participants (Number of Errors)			
Inexperienced	**5.82**[*] (3.76)	**4.89**[*] (5.36)	5.37 (4.59)
Experienced	**2.47**[*] (2.53)	**1.50**[*] (3.41)	1.97 (3.01)
Overall	4.72 (3.73)	3.68 (4.99)	4.21 (4.41)
Panel D: Overconfidence (Underconfidence) of Participants[b]			
Inexperienced	**1.67**[*],[***] (2.90)	**2.53**[*],[***] (2.08)	**2.08**[***] (3.56)
Experienced	**0.26**[*] (2.31)	**0.70**[*] (3.03)	0.49 (2.67)
Overall – H3	**1.21**[***] (2.78)	**1.88**[***] (1.54)	**1.54**[***] (3.36)

[a]All participants answered questions regarding process confidence in tax return preparation under both aided and unaided conditions.
[b]Measured as the miscalibration of participants, or the difference between actual errors and self-perceived errors in the tax return preparation task.

[*]Significantly different between inexperienced and experienced groups at $p < 0.05$
[**]Significantly different between aided (with TDSS) and unaided (without TDSS) conditions at $p < 0.05$
[***] Level of miscalibration of the participants (the difference between actual and perceived errors) is significantly different from zero (i.e., perfect calibration) at $p < 0.001$. Positive (negative) values represent overconfidence (underconfidence).

Table 4
Influence of *Ex-Ante* Confidence on Performance Accuracy (H1)

	Unaided Group – Number of Errors[a]	**Unaided Group – Dollar Value of Errors[a]**	**Aided Group – Number of Errors[a]**	**Aided Group – Dollar Value of Errors[a]**
Constant	**6.267***** (1.246)	**12102***** (3582)	**11.454***** (2.842)	**4221***** (1238)
Ex-Ante Confidence (H1)	.012 (.022)	-90.83 (64.36)	**-.072**** (.036)	**-33.88**** (15.5)
Number of Returns Prepared	**-.933***** (.334)	-742.04 (961.5)	-.420 (.397)	-11.78 (173.1)
Tax Work Experience	-.052 (.045)	24.88 (128.2)	-.040 (.069)	-8.742 (30.13)
Taxation Classes	-.396 (.461)	-710.77 (1325)	-.696 (.755)	-285.9 (328.8)
Familiarity with Software			-.036 (.368)	-2.859 (160.3)
Adjusted R-Squared F-Statistic	.193 **4.40*****	.041 1.610	.142 **2.817****	.063 1.735

[a]Coefficients (Standard Errors) for OLS Regression Models, by Dependent Variable

*, $p < 0.10$; **, $p < 0.05$; ***, $p < 0.01$; two-tailed tests, except for H1, which is directional (one-tailed)

Notes: For the OLS regressions for the unaided group, *ex-ante confidence* refers to participants' confidence in their ability to accurately prepare a tax return without tax preparation software (manually). For the OLS regressions for the aided group, *ex-ante confidence* refers to participants' confidence in their ability to accurately prepare a tax return with tax preparation software (with a TDSS).

Table 5
Influence of Experience on *Ex-Ante* Confidence (H2)

		F-Statistic	Significance Level (One-Tailed)
Panel A: ANOVA, *Ex-Ante* Confidence without TDSS (manually)			
Experience Level (H2)		13.232	**<0.001**
Adjusted R^2 = .098			
Panel B: ANOVA, *Ex-Ante* Confidence with TDSS (tax preparation software)			
Experience Level (H2)	1.288	0.259	
Adjusted R^2 = .003			

Panel C: Regression Analyses		
	Ex-Ante Confidence without TDSS (Manually) [a]	**Ex-Ante Confidence with TDSS (Software)** [a]
Constant	**45.00***** (4.04)	**67.24***** (4.05)
Number of Returns Prepared (H2)	**6.446***** (1.304)	**3.131***** (1.235)
Tax Work Experience	.137 (.215)	.246 (.192)
Taxation Classes	1.90 (2.23)	-.361 (1.99)
Familiarity with Software		**2.59**** (1.03)
Adjusted R-Squared F-Statistic	.229 **12.171*****	.166 **6.617*****

[a]Coefficients (Standard Errors) for OLS Regression Models, by Dependent Variable

*, $p < 0.10$; **, $p < 0.05$; ***, $p < 0.01$; two-tailed tests, except for H2, which is directional (one-tailed)

Table 6
Factors Influencing Overconfidence (H4 and H5)

	F-Statistic	Significance Level*	
Panel A: ANOVA, All Participants			
Model	2.483	0.065	
Experience Level (H4)	6.167	**<0.01**	
Aided/Unaided Condition (H5)	.995	0.321	
Experience * Aid Interaction	.106	0.745	
Adjusted R^2 = .038			
Panel B: ANOVA, Participants with Errors			
Aided/Unaided Condition (H5)		16.807	**<0.001**
Adjusted R^2 = .189			

*Significance level is one-tailed for directional tests (H4) and two-tailed for non-directional tests (H5)

Table 7
Examples of Common Errors for Participants <u>with Errors</u>

Type of Error	% Error – <u>Without</u> TDSS (Unaided)	% Error – <u>With</u> TDSS (Aided)
Dependents – Omitted from return	0%**	13%**
Wages	7%**	26%**
Child Tax Credit	76%***	22%***
Schedule A – Sales Tax	15%***	52%***
Schedule A – Real Estate Tax	11%***	52%***
Schedule A – Mortgage Interest	13%*	30%*
Schedule A – Charitable Contribution	13%**	35%**
Schedule C-EZ – Business Expenses	43%**	74%**

*, $p < 0.10$; **, $p < 0.05$; ***, $p < 0.01$; two-tailed tests