4-17-2014

To Trust or Not: The Effects of Monitoring Intensity on Discretionary Effort, Honesty, and Problem Solving Ability

Brian K. Laird

TO TRUST OR NOT: THE EFFECTS OF MONITORING INTENSITY ON DISCRETIONARY EFFORT, HONESTY, AND PROBLEM SOLVING ABILITY

by

Brian K. Laird

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: Business Administration with an emphasis in Accounting

The University of Memphis

May 2014

ACKNOWLEDGMENTS

I would like to thank my wife Jaime Chipman Laird and my daughter Jaida Michelle Laird for their encouragement and patience during my graduate studies. Without their support I could not have endured. I thank my advisor Dr. Charles Bailey for tirelessly helping me form, format, and polish my dissertation work. I am grateful to the University of Memphis and the Fogelman College of Business for providing financial support during my graduate coursework. Also, I thank the Institute of Management Accountants for providing funding support toward my dissertation experiments through their Doctoral Student Research Grant Program. Finally, I appreciate the faculty and staff at the University of Memphis and Arkansas State University for their support during all phases of my higher education. May God bless you all for the great work that you do!

ABSTRACT

Laird, Brian K., Ph.D. The University of Memphis. May 2014. To Trust Or Not: The Effects of Monitoring Intensity on Discretionary Effort, Honesty, and Problem Solving Ability. Major Professor: Charles D. Bailey

Managerial accounting researchers and practitioners are increasingly concerned with the effects of formal organizational controls on agent behavior. This three-paper dissertation extends this line of research by experimentally examining the effects of monitoring intensity on three important work behaviors which, generally, are not directly observable by the organizational control system: discretionary effort, problem solving ability, and honesty. Together, these studies help fill a gap in the managerial accounting literature by examining the relationship between the monitoring environment and agent behavior.

The principal-agent theory of the firm suggests that tighter monitoring by the principal will increase the agent's work effort at best, and have no effect at worst. However, the psychology literature suggests that monitoring may actually reduce effort by "crowding out" an individual's intrinsic motivation to perform unmeasured or unrewarded work related tasks. In Paper 1, I test for the crowding out effect of monitoring and find mixed results.

In Paper 2, I investigate the effects of monitoring intensity on various aspects of problem solving ability and creativity. Past research suggests that strict environmental controls can have detrimental effects on creative thinking. I extend this line of literature by investigating how monitoring affects an individual's problem solving ability. In general, I find that monitoring intensity is negatively associated with problem solving ability.

In Paper 3, I investigate how monitoring intensity affects an individual's propensity toward dishonesty using a 3x2 experimental design where the participants are given a simple task, with a monetary reward based on performance, in one of the three monitoring treatments—trust, human monitoring, or electronic monitoring—and in one of two outcome reporting regimes—self-report or verified. I find an inverted-U shape relationship between monitoring intensity and dishonesty, where dishonesty is highest under human monitoring.

Organizations are increasing their use of all types of surveillance and controls, and, in general, trust is increasingly discouraged within organizations. These papers add to the managerial accounting literature by shedding light on how different monitoring environments can change human behavior. This line of research can only increase in importance as regulation increases and monitoring technology becomes more advanced, reliable, and accessible.

PREFACE

This three-paper dissertation experimentally examines the effects of monitoring intensity on three important work behaviors which are, generally, unobservable by the organizational control system: discretionary effort, problem solving, and honesty. The three studies are written as independent chapters for potential publication, each of which has been submitted to the American Accounting Association's Annual Meeting and other academic conferences. As a result there may be repetition of some information in the studies.

The three experiments that comprise this dissertation were done in one sitting for each participant. Twelve sessions were held in a computer lab at a large public university. Each session contained either 9 or 10 participants. A total of 114 individuals participated, earning an average of $15.15. The total payout to participants was $1,728.

At the end of this dissertation there is a general conclusion. The general conclusion ties the three studies together and discusses the implications of the dissertation as a whole. The implications of each study also are discussed in a conclusion section at the end of each study.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

This three-paper dissertation experimentally examines the effects of monitoring intensity on three important work behaviors which, generally, are not directly observable by the organizational control system: discretionary effort, problem solving, and honesty. The basis of this research is that most individuals, by default, are internally motivated to exert effort in order to perform a fair exchange, be honest, and utilize complex problem solving skills. However, control mechanisms can crowd out the internal motivation to be fair and honest by lowering the individual's propensity toward reciprocity (Gächter and Falk 2002) or by lowering the individual's threshold for dishonesty.[1] Additionally, controls may increase work-related stress which can negatively influence problem solving skills (e.g., Hennesey and Amabile 1998; Elsbach and Hargadon 2006). These effects may be more pronounced when the control system is perceived by the agent as being intrusive, overly controlling, or unnecessary (Stanton 2000).

Often the terms monitoring and control are used interchangeably in the business literature. However, most formal definitions of the two terms view monitoring as one part of the control system. Tosi et al. (1997, 588) defined monitoring as "observation of an agent's effort or outcomes that is accomplished through supervision, accounting controls, and other devices." Monitoring, alone, is void of any rewards, punishments, or corrective actions. Koontz and O'Donnell (1955, 103), in the classic book *Principles of Management,* describe control as "the measurement and correction of performance in order to make sure that enterprise objectives and the plans devised to attain them are

[1] Research has shown that individuals' preference for honesty often depends on environmental factors (e.g., Evans et al. 2001; Belot and Schröder 2013).

accomplished." Through monitoring, effort and/or outcomes are observed and measured. Monitoring becomes part of the organizational control system when these observations and measurements are used to influence future performance. Thus, the control system can be broken up into two parts, monitoring (measurement or observation) and control (corrective actions, rewards, threats, compensation scheme, etc.), and each part can be studied separately. In these studies, I consider the effects of monitoring intensity on certain aspects of behavior while holding the other parts of the control system constant.

The three experiments that comprise this dissertation were done in one sitting for each participant. Twelve sessions were held in a computer lab at a large public university. Each session contained either 9 or 10 participants, with a total of 114 individuals participating. Each session was pre-assigned one of three monitoring treatments, trust monitoring (low), human monitoring (medium), or electronic monitoring (high), and separately, one of two reporting treatments, self-report or verified, for the Task 3. Each participant was assigned to one treatment group and performed three tasks, with each task representing a new experiment.

My first study looks at the effects on monitoring intensity on discretionary effort. The principal-agent theory of the firm suggests that tighter monitoring by the principal will increase the agent's work effort at best, and have no effect at worst. However, standard principal-agent theory doesn't consider the effects of monitoring on agent work behaviors that falls outside of the control system, such as discretionary effort or voluntary effort. The psychology literature suggests that monitoring may actually reduce such effort by "crowding out" an individual's intrinsic motivation to perform unmeasured or unrewarded work. Accordingly, I hypothesize that as monitoring increases across groups,

discretionary effort and voluntary effort will decrease. The details of the study, results, and implications are discussed in Chapter 2.

In the second study I investigate the effects of monitoring intensity on various aspects of problem solving ability and creativity. Past research suggests that strict environmental controls can have detrimental effects on creative thinking, which is critical for finding the optimal solution to complex problems. I test this proposition and I extend this line of literature by investigating how monitoring affects an individual's ability to establish and use a pattern solution, recognize when a pattern solution is no longer efficient, and solve complex problems. The details of the study, results, and implications are discussed in Chapter 3.

In the third study I investigate how monitoring intensity affects an individual's propensity toward behavioral dishonesty. This variable is operationalized by splitting each monitoring treatment in to two subgroups. Each group is given the same simple puzzle, with monetary compensation tied to performance. One group self-reported their performance, while the other group had their performance checked.[2] Past research has shown that individuals are very prone to reciprocal behavior, with such behavior extending into the principal-agent context. I hypothesize that as monitoring increases across groups, dishonesty will also increase. The details of the study, results, and implications are discussed in Chapter 4.

Organizations are increasing their use of all types of surveillance and controls, and, in general, trust is increasingly discouraged within organizations. These papers add to the managerial accounting literature by shedding light on how individuals react to

[2] This research design, which test honesty, is similar to the research design used in Ariely et al. (2009).

various levels of monitoring and extend a larger body of research that is concerned with the effects of formal controls on behavior within organizations (e.g., Christ et al. 2012). Specifically, this research is concerned with the potential "hidden costs of monitoring" which may arise during the organizational control process. This line of research can only increase in importance as regulation increases and monitoring technology becomes more, advanced, reliable, and accessible.

Because of the complexity of the individual psyche and the vast number of uncontrollable factors in organizational settings, investigations into the effects of monitoring on individual behavior are well-suited for laboratory experiments. The results of these experiments are not meant to be directly generalizable outside of a laboratory setting. Rather, these experiments, and their results, should be considered in the larger frameworks of organizational theory, human psychology, and current practices.

CHAPTER 2

THE EFFECTS OF MONITORING INTENSITY ON DISCRETIONARY TASK EFFORT, VOLUNTEER RATES FOR OPTIONAL TASKS, AND EFFORT ON OPTIONAL TASKS

INTRODUCTION

The Principal-Agent theory of the firm suggests that tighter monitoring by the principal will increase the agent's work effort (Alchian and Demsetz 1972) at best, and have no effect at worst. In contrast, the psychology literature suggests that monitoring may actually reduce effort by "crowding out" an individual's intrinsic motivation to perform a task (Frey 1994) or to be "fair" to the principal (Gächter and Falk 2002). The agency theory and the "crowding out" literature are not necessarily contradictory if one considers that most principal-agent relationships involve some effort that is monitored by the principal, and some effort that cannot be, or is not, monitored (Hölmstrom 1979; Bernheim and Whinston 1998; Hecht et al. 2012). Building on these ideas, I attempt to reconcile the agency theory literature and the human motivation literature by distinguishing between the monitored aspects of a task, in which external controls may substitute for internal motivation, and the less monitored discretionary and optional aspects of a task, in which external controls may not substitute for internal motivation. I hypothesize that increasing task-monitoring (external control) will increase effort on the monitored aspects of a task at the expense of discretionary or voluntary aspects of the task. Accordingly, this paper fills a gap in the managerial literature by experimentally examining the relationship between task-monitoring intensity and three different work behaviors generally associated with internal motivation: discretionary effort on a

mandatory task, volunteer rates for a work-related optional task, and effort exerted on the optional task.

I used a laboratory experiment where the participants, 114 in total, were assigned to one of three different monitoring treatment groups, electronic monitoring, human monitoring, or trust (no monitoring). They were then asked to perform the clerical task of cross-checking paper invoices with error-riddled transcriptions of the data, which had been pre-entered into a spreadsheet. Each treatment group performed the clerical task for a flat wage. The only task instruction for the participants was that they must correct records for the entire work period to receive the flat wage. No direct instructions pertaining to work quality or work quantity were given, leaving both to the discretion of the worker. However, quantity was more observable by the monitoring than quality, making quality more discretionary. After the task, and during a short free-time break, the participants were solicited for optional feedback. On the clerical task, work task quantity and work task quality (representing discretionary effort) were measured, while on the optional feedback task, volunteer rates for the optional task and task completion rates (representing effort spent on the optional task) were measured.

As discussed by Frey (1993), it is possible that, under certain conditions, increasing monitoring may actually reduce agent effort by simultaneously lowering the marginal cost of shirking and increasing the marginal cost of effort to the agent. Such conditions abound when the monitoring is imperfect, providing an opportunity to decrease effort, and the agent is psychologically affected by the monitoring or the change in relationship with the principal, prompting a decrease in motivation. In this research, I hypothesize that the more monitored an individual feels, the less obligated, and less

motivated, the individual will feel to perform tasks that are not directly monitored or measured by the control system, as compared to individuals who do not feel intensely monitored. I find evidence to support this hypothesis when examining monitoring intensity and discretionary effort (work quality) on a mandatory task. In this experiment, work quality was higher when monitoring intensity was low. However, I find a more complex relationship between monitoring intensity and optional tasks. Contrary to my hypothesis, I find evidence, through higher volunteer rates, that monitoring intensity is positively associated with the propensity to work on tasks that are presented as optional. However, I also find that monitoring intensity is negatively related to effort spent on the "optional" task once an individual volunteers to do it. This paradox likely manifests because individuals, having been monitored in the previous task, still feel controlled. This feeling of being controlled compels them to volunteer for the optional task, but they have less internal motivation to sustain effort toward, or complete, the optional task.

Most work-related tasks are not intrinsically motivated, but are to some degree externally motivated. Activities that are not intrinsically motivating require extrinsic motivation, so their initial enactment depends upon the perception of a contingency between the behavior and a desired consequence. However, when certain conditions are met, individuals will adopt the actions, or goals, that were initially externally motivated as part of their autonomous behavior so that the external contingency is no longer needed to invoke the performance. Self-determination theory (Gagné and Deci 2005) postulates that the more an individual feels controlled, the less likely they are to internalize external regulations, values, and rules. This theory could account for monitoring intensity being

positively related to volunteer rates on "optional" tasks, yet negatively related to performance on the "optional" tasks.

This paper adds to the prior literature by showing evidence that employee monitoring, as an organizational control mechanism, has negative effects on internal motivation as compared to trust and reciprocity. However, the effects of lower internal motivation are not readily visible on the monitored aspects of tasks, but are more likely to manifest around the less monitored aspects of the task or the parts of the task which are perceived by the agent as being optional. These findings suggest that when the principal-agent relationship involves a complex job design, it may be more beneficial for the principal to rely on trust and reciprocity instead of control, especially if the principal can only monitor, or measure, part of the agent's overall job or output.

The rest of this paper is organized as follows: Section II provides more theory and background, Section III develops the hypotheses, Section IV describes the research design, Section V provides the analysis of the results, and Section VI discusses extensions, limitations, implications, and conclusions of the study.

THEORY AND BACKGROUND

Monitoring and Agency theory

In accounting and finance-related research, agency theory dominates the discussion of organization control and management. Lambert (2001, 3) states that "agency theory has been one of the most important theoretical paradigms in accounting during the last 20 years. The primary feature of agency theory that has made it attractive to accounting researchers is that it allows us to explicitly incorporate conflicts of interest, incentive problems, and mechanisms for controlling incentive problems into our models."

In the standard agency theory model, income earned provides benefits while the effort spent to earn it is a disutility. Further, agents will always exploit opportunities to lower their effort absent a penalty for doing so. The introduction of more monitoring cannot lower effort since agents are naturally effort averse (Alchian and Demsetz 1972) and are, presumably, already maximizing income and minimizing effort, on average, at any given point in time.

Despite its successes, agency theory has been criticized for the diminished realism of adhering strictly to narrow self-interest and ignoring nonmonetary preferences such as ethics, trust, and fairness (Arrow 1985). Many modern corporate contracts, control systems, and governance structures are designed and based on the principles of agency theory. This reality essentially means that many of the corporate "best practices" accepted today do not emphasize important psychological components. Arce (2007) examined how the assumptions of agency theory may be self-activating. He did so by exploring a principal-agent framework that allows for the possibility that rational agents may hold intrinsic preferences for autonomy in decision making and experience disutility from being monitored. His analysis identified conditions under which the economic approach to agency, which is principally framed in terms of monetary rewards and the avoidance of effort, can select against agents' intrinsic preferences for autonomy and break implicit contracts between principal and agent that are based on trust. In short, if controls are built purely on economic rationality, then that is the type of behavior they will cultivate. For example, agency-theory-based contracts have been accused of encouraging opportunistic behavior (Ferraro et al. 2005) and blamed for the deteriorating moral climate that has given rise to Enron and other corporate scandals (Kulik 2005).

This experiment builds upon a long and interesting line of research that attempts to add a human psychological component to the standard agency theory assumptions. I test the standard agency theory assumption that monitoring cannot decrease effort, by examining how monitoring affects performance on the aspects of tasks that the agent may consider discretionary or voluntary. The goal of this line of research is to add descriptive, prescriptive, and pedagogical usefulness to the principal-agent model (Stevens and Thevaranjan 2010).

Monitoring and Effort

In a widely cited discussion article, Frey (1993) asked the rhetorical question "does monitoring always increase effort?" Frey's main concern was the triangular connection of monitoring, trust, and effort in the principal-agent relationship. He concluded that the effects of monitoring on effort depend on whether the agent perceives the monitoring to be a signal of distrust, which is more likely to be the case when the agent and principal have an inter-personal relationship. When a psychological contract exists between the agent and principal, an increase in, or focus on, monitoring may be seen as violation of the mutual trust that has been established in the relationship. With their previous psychological bond broken, the agent now has a lower marginal benefit from working and higher marginal benefit from shirking. Shirking is now more likely, assuming that the agent has the opportunity.

Dickinson and Villeval (2008) tested Frey's (1993) theory in a laboratory setting. They were interested in how anonymous versus interpersonal auditing would affect effort. Their treatments were applied by having a portion of the participants meet with the individual serving as their monitor and sit by their monitor during the experiment, while

the other participants never met or saw their monitor. The task involved using a computer to move electronically along a line for a performance contingent wage. Each move cost the operator monetarily (a proxy for effort), and, the monitor's payoff depended on the operator's performance. If the operator was audited and underperformed, a fine was incurred by the operator and retained by the monitor. They find that most agents react to high levels of monitoring by increasing performance. However, they find that above a certain threshold, monitoring decreases effort, and this effect is most pronounced in the interpersonal treatment. The current research design differs from Dickinson and Villeval (2008) because I set out to capture changes in discretionary effort and volunteer rates for optional tasks, and I am using actual work effort instead of a proxy.

In another interesting study, Callahan and Larson (1990) tested whether performance monitoring can influence work behavior in the absence of any subsequent managerial action or feedback. They postulated that monitoring activity alone can serve as cue, signaling the relative importance of one task over another. In their experiment, they gave each participant two tasks, which they were to work on concurrently for two hours. In the control group, the participants were left alone for two hours. In the treatment group, the experimenter would come in every 20 minutes and check the progress of one of the tasks only, offering no feedback. Despite being instructed to work diligently on both tasks, the treated participants outperformed on the monitored task while performance on the non-monitored task fell in proportion. The net effect of monitoring, when total production on both tasks was considered, was small as compared to a control group who worked on both tasks without any monitoring.

Falk and Kosfeld (2006) looked at the effects of control on effort. They found that control entails a hidden cost as most participants (agents) reduced their effort, which was chosen at a cost to the agent, in response to hurdles set by the principal. When asked for their emotional perception of control, most agents who reacted negatively said that they perceived the controlling decision, to set a hurdle, as a signal of distrust and a limitation of their autonomy. Falk and Kosfield (2006) suggest that agents do not like being restricted. They also suggest that agents perceive controls as a signal of distrust and low expectations.

Employee reactions to monitoring and control matter because organizations have a strong stake in maintaining both employee motivation and well-being. Most researchers suggest that monitoring technology itself is neutral, and that it is the design and implementation of the technology that affects employee reactions (Alder and Ambrose 2005). For instance, Stanton and Weiss (2000) claim that when employers provide adequate justification for monitoring there are generally few negative effects. However, most of this evidence comes from employee surveys and interviews, as there is very little empirical evidence on the matter.

HYPOTHESIS DEVELOPMENT

The negative effects of rewards, incentives, and punishment on internal motivation have been well documented by researchers (Gneezy et al. 2011; Deci et al. 1999). Behavioral theorists predict that the same phenomenon, known as "crowding out," will also be a factor in the relationship of monitoring and internal motivation (Frey 1994). Because monitoring has a direct effect on task effort through the control system, studying the links between monitoring, intrinsic motivation, and effort is difficult. In the past,

researchers have overcome this problem by using intrinsically motivating tasks, such as puzzle solving. Researchers have manipulated the control system, and then measured how well the participants performed on the task, their attitude toward the task, or the time they spent working on the task during free-time (Wiersma 2011). However, the value of this research, and its theories, is limited for managerial research because most work-related tasks are not purely intrinsically motivating, yet people still perform work-related tasks with varying levels of performance under various levels of controls and monitoring. For this reason, Self-determination theory (Gagné and Deci 2005) serves as more practical way of viewing motivation for work-related tasks.

Self-determination theory posits that motivation represents a continuum from no motivation (amotivation) to completely internal motivation, with different levels of external motivation in the middle. The levels of external motivation range from being completely controlled and performing a task, to being completely autonomous and performing a task. One major point is that when one autonomously performs a task, it does not necessarily mean that the task is intrinsically motivating. Self-determination theory suggests that intrinsic motivation concerns experiencing activities as being interesting and spontaneously satisfying, whereas autonomous extrinsic motivation concerns experiencing activities, not as interesting or fun, but as personally important for one's self-selected aims, goals, and purposes (Gagné and Deci 2005). Autonomous extrinsic motivation results from the internalization of an extrinsically motivated behavior into a personally endorsed behavior. The importance of this theory, especially to managerial research, is that it shows how monitoring and control can affect motivation on tasks that are not inherently intrinsically motivating. According to the Self-determination

theory, the monitoring and control of an agent should increase external motivation and decrease the internalization of any goals or values associated with the task or job. Conversely, trust and autonomy should increase internal motivation and increase the internalization of task goals. In my hypothesis, I attempt to test these propositions.

Mandatory Task Quantity and Quality

When the agent is performing a simple mandatory task, monitoring may be a substitute for internal motivation. Consistent with traditional agency theory, since most individuals are not intrinsically motivated to perform common work tasks, monitoring should increase effort spent by the agent. However, the increase effort may be focused only on the monitored aspects of the task since the agent has less internal motivation to focus on other aspects of the task.

If monitoring increases external motivation and lowers internal motivation, then it should negatively affect discretionary effort, as shown in Figure 2.1. In this experiment, monitoring of the agent is such that quality is less observable, more difficult to measure, and left more to the discretion of the worker, than quantity. In the event of a loss of intrinsic motivation that may result from monitoring, quality is likely to suffer before quantity.[1] This leads to the following substantive hypotheses about quantity and quality:

H1a: All else equal, production *quantity* increases as the intensity of monitoring increases.

H1b: All else equal, production *quality* decreases as the intensity of monitoring increases.

[1] Hypothesis 1b may fail to be supported if monitoring drives quantity, and there is a high level of performance spillover (Hecht et al. 2012), where an increased attention to one task measure (quantity) positively affects other areas of the task (quality).

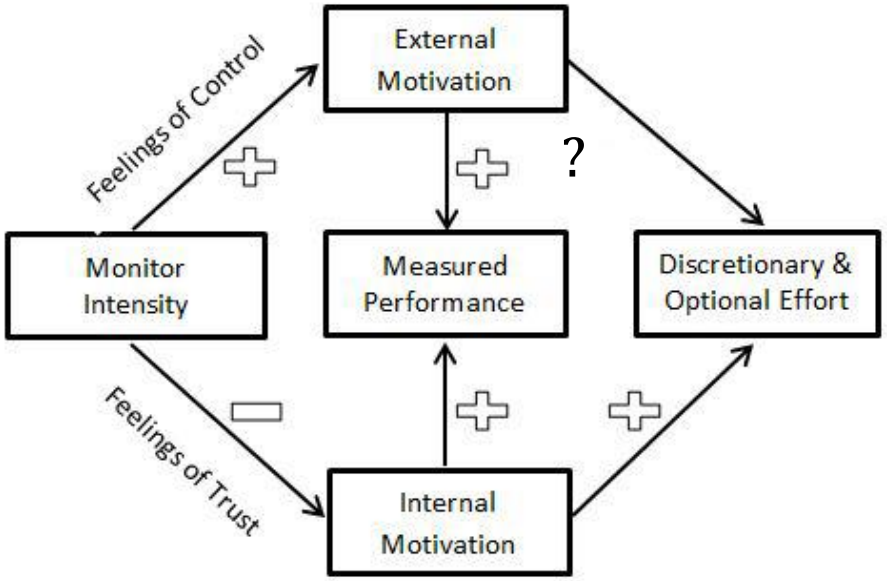


Figure 2.1.Thoeretical Model: The link between external motivation and discretionary or optional effort is weak and may depend on the agent's perception that it truly is optional.

Optional Task Volunteer Rates and Effort

Some researchers have suggested that performance and effort, on all but the simplest job design, can be categorized into two parts, task performance and contextual performance (see Motowildo and Van Scotter 1994). Borman and Motowidlo (1993) defined task performance as activities that are formally recognized as part of the job and that contribute to the organization's technical core, while they defined contextual performance as individual behavior that is discretionary and that in the aggregate promotes the effective functioning of the organization. For example, when an individual performs extra work tasks voluntarily, or provides thoughtful ideas and feedback to management, that would qualify as contextual performance but not necessarily as task

performance. While it may be difficult to disentangle the effects of "internal motivation versus control" on the various aspects of mandatory task performance, it should not be as difficult to disentangle the effects of "internal motivation versus control" on contextual performance because contextual performance often falls outside of the control system, and should be driven, mostly, by internal motivation. If monitoring decreases intrinsic motivation, contextual performance should decrease as well. From this logic I draw my last two hypotheses.

H2a: All else equal, increased task-monitoring intensity decreases volunteer rates on optional tasks.

H2b: All else equal, increased task-monitoring intensity decreases the effort spent on optional tasks.

RESEARCH DESIGN

This experiment was carried out in a computer lab at large public university. Twelve sessions were held and each session included either 9 or 10 participants. A diverse group of 114 adult volunteers participated. Participants were recruited through the university email newsletter, flyers, and word of mouth. Participants self-registered online and chose the session date and time they preferred. Upon arrival, the participants were given a short demographic survey, as shown in Appendix A. Question 6 was a distractor to help disguise the purpose of the study, and 7 through 9 were exploratory, potential covariates. Table 2.1 shows the key demographics collected from the participants.

Table 2.1
Demographics of Research Sample by Treatment Group

Gender	Trust Monitored	Human Monitored	Electronically Monitored	Total	Percent
Male	19	17	18	54	47%
Female	19	21	20	60	53%
				114	100%
Age					
18-24	23	26	18	67	59%
Over 24	15	12	20	47	41%
				114	100%
Student Nationality					
Domestic	23	24	21	68	60%
International	14	10	12	36	32%
Not a Student	1	4	5	10	9%
				114	100%
Business Student					
Business Student	14	11	11	36	32%
Non-Business Student	23	23	21	67	59%
Not a Student	1	4	6	11	10%
				114	100%
College Level					
Fresh/Soph	19	18	8	45	39%
Junior/Senior	13	8	12	33	29%
Graduate	5	8	10	23	20%
Non Student	1	4	8	13	11%
				114	100%

The sessions were pre-assigned one of three monitoring treatments: electronic monitoring as the most intense treatment,[2] human monitoring as a mid-level treatment, and trust (no monitoring) as a low-level treatment. When the participants arrived they were provided a consent form and demographic survey. Next, their attention was directed to the white board where the task schedule and compensation plan were explained. All the participants were told that they were being paid a $10 flat wage to complete 27-minute task. They were told that the flat wage meant that it did not matter how much work they performed but they must work the entire time. The participants were told that they would have a short free time break after the task before moving on to different tasks. The short break allowed for the chance to offer optional work, as explained below.

Individuals in the electronic monitoring treatment had one small webcam facing their workstation keyboard and papers when they arrived, as shown in Appendix B. The goal of the webcam placement was to create the perception that the focus of the monitoring was on quantity of work. This monitoring design is tantamount to input monitoring (see Pendergrast 2000) where the agent's inputs, such as work time, progress, and resources are closely watched, but the output is unknown or unmonitored. The purpose of this monitoring design was to create a task control system where work quality was more discretionary than work quantity. Participants in the electronically monitored group received all the same task instructions as the other two groups, except they were told that "you are being monitored with webcams so we can observe your work and make

[2] While electronic monitoring and traditional human monitoring have the same fundamental purpose, past research suggest that the pervasive, continuous nature of electronic monitoring often elicits stronger reactions from the worker (Aiello and Kolb 1995; Lund 1992; Stanton 2000).

sure you follow the instructions as given."[3] Unbeknownst to the participants, the webcams were not activated. However, the experimenter and research assistant sat prominently at a corner workstation that the subjects believed to be the "monitoring station." The setup was designed so that when the monitoring station was manned the subjects believed they were being monitored and when the station was unmanned the subjects believed they were not being monitored. The station was manned for the entire task period but not during the free time break when the optional task was offered.

Participants in the human monitored group were subjected to traditional human monitoring and received all the same task instructions as the other two groups, except they were told that "I will walk around the room so I can observe your work and make sure you follow the instructions as given." The researcher and the research assistant wandered around the room and passively observed the participants during the task time but left the room during the free time break when the optional task was offered. .

Individuals in the trust treatment received all the same task instructions as the other groups. However, they were told that "you will not be watched and we believe you will follow instructions as given." In this treatment, all research personnel then left the room and returned when time was up[4] and then left again during the free time break when the optional task was offered.

All the participants performed a data correction task, based on the experimental design used in Stanton and Sarkar-Barney (2003). The computer at the participant's

[3] This wording used in the participant instructions is based on the wording used in the study by Enzle and Anderson (1993). Their study looked at the effects of controlling versus non-controlling electronic monitoring on intrinsic motivation.

[4] Rousseau et al. (1998, 395) defined trust as "a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behavior of another."

workstation had a spreadsheet opened up. Each row on the spreadsheet represented a different invoice, and each column contained a data point such as invoice number, invoice date, customer name, customer address, items bought, total due, etc. There were a total of 24 data points for every one of the 50 invoices provided. Each participant was told to crosscheck the electronic database records against hardcopies of the invoices, which were bound in a folder at the workstation. The participants were told they were checking behind an individual who entered the data very quickly and that they each had different invoices except for the first sample invoice; all actually received the same data. The sample invoice was worked with them, on a projector, by the researcher as an example. Each invoice had between one and five errors, randomly distributed, with an average of three errors per record. Appendix C shows an actual invoice used, and Appendix D shows part of the spreadsheet used in the experiment.

After the participants completed the 27-minute task, they were told that they were to take a short "free time" break where they could rest, check their cell phone, or open up the Internet browser.[5] They were asked not to speak to anyone in person or on the phone. Immediately after being informed of the free time options, they were told the following: "Also, let me direct your attention to the back of the invoice folder. There you will find three feedback forms. The feedback you provide helps us to improve the task you performed. The feedback forms are optional and not required." All research personnel then left the room for the remainder of the break. The participants had approximately 5 minutes of break time. The three feedback forms requested responses on a Likert Scale

[5] This was the first of three experiments for which the participants had volunteered that day, providing the context for this "break" time. The participants were told in the consent form that they should not leave the room until the session is completely over.

had eight questions per form, and could be filled out in less than 1 minute per form, as shown is Appendix E.

Two measures were taken from the analysis of this optional task: whether the participant volunteered (binary), and if so, the effort expended. If the participant filled out any forms (volunteered) then effort expended was measured by the participant's outcome on the optional task. If the participant only partially completed the voluntary optional task by filling out 1 or 2 feedback forms than effort expended was lower than if the participant completed the task by filling out all 3 feedback forms.

RESULTS

Mandatory Task Quantity

Consistent with traditional agency theory**,** Hypothesis 1a predicts that, all else equal, production *quantity would* increase as the intensity of monitoring increased. The mean number of invoices checked was 17.16 for the trust treatment, 17.29 for the human monitored treatment, and 16.68 for the electronically monitored group, as shown in Panel A of Table 2.2. An ANOVA (Panel B of Table 2.2) shows no significant difference (p= .83) between treatment groups in the quantity of work performed on the assigned task. Including covariates to the analysis did not change this result. The power of the ANOVA is approximately .60, assuming a medium effect size and alpha risk set at .10. Therefore, the hypothesis is not supported. The monitoring regime had little effect on production task quantity.

Table 2.2
Summary Statistics and ANOVA for Work Quantity by Monitoring Treatment

Panel A: Summary Statistics

Treatment	Obs.	Mean	Std. Dev	Min	Max
Trust	38	17.16	4.67	10	26
Human Monitoring	38	17.29	4.63	6	29
Electronic Monitoring	38	16.68	4.54	8	28
Total	114	17.04	4.58	6	29

Panel B: ANOVA Results

Number of Observations	114	**R- squared**	0.00
Root MSE	4.61	**Adj R-squared**	-0.01

Source	Partial SS	df	MS	F	Prob > F
Model	7.70	2	3.85	0.18	0.83
Treatment	7.70	2	3.85	0.18	0.83
Residual	2361.08	111	21.27		
Total	2368.78	113	20.96		

Note: Work Quantity is measured as the number of records checked during the task time, regardless of quality. Individuals checked a box for each record they reviewed. This was cross-checked against the number of electronic records altered. No difference was found between the two measures.

Production Task Quality

Hypothesis 1b predicts that, all else equal, production *quality* will decrease as the intensity of monitoring increased. When monitoring was applied, the quality of the work was more discretionary than the quantity of work. For this reason, quality is considered discretionary effort. Quality was operationalized at the number of errors corrected divided by the number of errors in the invoices checked. The seeded errors ranged from one to five per document, and averaged around three errors per record regardless of the number of records that the participant completed.

The mean quality was .864 for the trust treatment, .816 for the human monitored treatment, and .820 for the electronically monitored group, as shown in Panel A of Table 2.3. The ANOVA in Panel B of Table 2.3 shows a difference in the quality of work performed on the task between treatment groups (p= .069). Including covariates to the analysis did not change this result. In a planned comparison between the Trust group and the two monitored groups, the Trust group shows a significantly higher rate of discretionary effort (p= .02), as shown in Panel C of Table 2.3. Thus, Hypothesis 1b is supported.

Table 2.3
Summary Statistics and ANOVA for Work Quality by Monitoring Treatment

Panel A: Summary Statistics

Treatment	Obs.	Mean	Std. Dev	Min	Max
Trust	38	0.864	0.083	0.578	0.970
Human Monitoring	38	0.816	0.109	0.455	0.952
Electronic Monitoring	38	0.820	0.107	0.474	0.978
Total	114	0.833	0.102	0.455	0.978

Panel B: ANOVA Results

Number of Observations	114	**R- squared**	0.05
Root MSE	0.100	**Adj R-squared**	0.03

Source	Partial SS	df	MS	F	Prob > F
Model	0.055	2	0.028	2.74	0.069
Treatment	0.055	2	0.028	2.74	0.069
Residual	1.121	111	0.010		
Total	1.176	113	0.100		

Panel C: Planned Comparison

	Contrast	Value of Contrast	Std Error	df	Sig. 2-tailed
Assuming Equal Variance	1	0.047	0.020	111	0.02
Not Assuming Equal Variance	1	0.047	0.018	93	0.01

Note: Work quality is measured as the number of errors corrected divided by the number of errors encountered during the task time, regardless of quantity. Errors were randomly distributed and averaged around three per record.

Volunteer Rates for Optional Task

Hypothesis 2a predicts that, all else equal, task-monitoring intensity will be negatively related to volunteer rates on optional tasks. Volunteer rates for optional tasks were measured by the willingness of the participants to offer any optional feedback during their free time break.

The proportions of volunteers were .737 for the trust treatment, .763 for the human monitored treatment, and .921 for the electronically monitored group, as shown in Panel A of Table 2.4. The results show, through a logistic regression, a significant difference in the volunteer rates for optional tasks across treatment groups. The electronically monitored group showed a higher propensity to volunteer ($p = .04$), as shown in Panel B of Table 2.4. The Chi-square analysis confirms a difference in the number of volunteers across treatment groups ($p = .09$).

Table 2.4
Optional Task Volunteer Rates by Monitoring Treatment: Summary Statistics, Logistic Regression, and Chi-Square Test

Panel A: Summary Statistics

Treatment	Obs.	Mean	Std. Dev	Min	Max
Trust	38	0.737	0.446	0	1
Human Monitoring	38	0.763	0.431	0	1
Electronic Monitoring	38	0.921	0.273	0	1
Total	114	0.807	0.396	0	1

Panel B: Logistic Regression

		Observations	114
Log likelhihood	-53.20	**LR chi2 (2)**	5.440
Pseudo R2	0.05	**Prob > chi2**	0.066

Volunteered	Coef.	Std. Error	Z	P>|z|
Electronic Monitoring	1.43	0.71	2.02	0.04
Human Monitored	0.14	0.53	0.26	0.79
Constant	1.03	0.37	2.79	0.01

Panel C: Chi-Square Test and Fisher's Exact

	Volunteered		
Treatment	**No**	**Yes**	**Total**
Trust	10	28	38
Human Monitored	9	29	38
Electronic Monitored	3	35	38
Total	22	92	114

Pearson's Chi-Squared Test

Pearson chi (4) = 4.8
Pr = 0.09

Fisher's Exact Test

Fisher's Exact = 0.09

Note: The paticipants had the opportunity to volunteer to provide feedback during a short free time break after their task. Volunteers were coded as 1 if the particpant volunteered to provide any optional feedback about the clerical task they performed for a flat wage, and 0 otherwise.

These results are contrary to Hypothesis 2a, which states that monitoring intensity will decrease volunteer rates for optional tasks. Conversely, I find evidence that the relationship may be positive. The analysis of Hypothesis 2b, concerning the *effort* expended, however, provides a more complete picture.

Effort Spent on Optional Tasks

Hypothesis 2b predicted that, all else equal, task-monitoring intensity would be negatively related to the effort spent on optional tasks by volunteers. In this experiment, the effort spent on optional tasks is measured by the participant's output on the optional task. If the participant only partially completed the voluntary optional task by filling out one or two feedback forms than effort expended was lower than if the participant completed the task by filling out all three forms. The results from the volunteer rate analysis, above, show that as monitoring increased more individuals volunteered for the optional task. Table 2.5 reveals, however, that as monitoring increases more people volunteer, but that they tend not to complete the task.

Table 2.5
Frequency Analysis of Optional Task Outcomes

	Outcome			
Treatment	**Abstained**	**Partial**	**Complete**	**Total**
Trust	10	9	19	38
Human Monitor	9	14	15	38
Electronic Monitor	3	19	16	38
Total	22	42	50	114

Stastistical Analysis

Pearson's Chi-Squared Test	**Fisher's Exact Test**
Pearson chi (4) = 8.0	**Fisher's Exact =** 0.08
Pr = 0.09	

Note: This table examines the effects of monitoring intensity on the outcomes of an optional task, filling out feedback forms about a mandatory task. Particiants either abstained, started but did not complete, or completed the task.

Once the decision was made to give feedback, the trust group spent the most effort on feedback. On average, the trust group filled out 2.43 feedback forms, the human monitored group filled out 2.14 feedback forms, and the electronically monitored group filled out 2.11 feedback forms (summary not shown). Since my measure of effort was the participant's outcome on the optional task, it is appropriate to use Pearson's Chi-square Test and Fisher's Exact Test, as shown in Table 2.5. The results show a significant statistical difference across groups between the number of individuals who abstain, give partial feedback, and give full feedback, at the .09 confidence level for the Pearson's Chi-square Test and .08 for the Fisher's Exact Test. The ratio of those of who start but fail to complete the optional task to those who start and finish the task is .47 to 1 for the trust group, .93 to 1 for the human monitored group, and 1.19 to 1 for the electronic

monitoring treatment. Overall the results support Hypothesis 2b, that all else equal, task-monitoring intensity is negatively related to the effort spent on optional tasks. Including covariates in the analysis did not change this result.

CONCLUSION

This study finds evidence that traditional agency theory models are wrong when they assume that monitoring cannot reduce worker effort. Further, this paper seems to indicate that there does not have to be a psychological bond broken between the agent and the principal for monitoring to have an effect on effort. Evidence is shown to support the idea that there are significant "hidden costs monitoring" which are often elusive to managers and researchers. The effects of monitoring often go unnoticed because they aggregate outside of the monitoring system, where lower internal motivation can be expressed with fewer repercussions for the agent.

These findings may also help to suggest, or explain, certain aspects of job design and performance such as those laid out by Holmstrom and Milgrom (1991). Holmstrom and Milgrom (1991) suggest that an employer, when faced with multiple employees and multiple jobs, should group jobs that can be easily measured together and assign those jobs to one group of employees, while assigning the other group of employees jobs that cannot be easily measured. This paper supports this recommendation, especially if the measurement process involves input monitoring.

It is reasonable to assume from the evidence shown in this paper that monitored individuals may work at a slightly lower effort levels than trusted individuals on mandatory work tasks. In this experiment, there were no difference in work quantity, but the trust monitored group had higher work quality. These finding support the conclusion

of past studies that show the movement toward more flexible workplaces (Shepard et al. 1996) and telecommuting increase overall output (Westfall 2004).

In this study, as monitoring increased so did volunteer rates, even though the monitor left the room during the time the optional task was performed. There are at least three reasons why this outcome may have occurred. First, it is possible that the monitored individuals assumed they were still being monitored even when the monitor had left the room. Second, the control of the monitoring may have carried over psychologically for a short time. Third, the controlled individuals may have cherished the return of their freedom. It is unknown if one these possibilities, or a combination, led to the higher volunteer rates. However, one thing is clear, as volunteer rates for the optional task increased, the average effort spent on the optional task, by those who volunteered, decreased. This decrease in effort is indicative of lower internal motivation.

Future research should focus more on learning how internal controls affect the performance of optional work tasks and discretionary parts of mandatory tasks. With the increasing complexity of the work environment and the increasing demand for customer service, optional and discretionary effort is increasingly important. Another interesting avenue of research is the interaction of pay schemes and monitoring intensity on optional and discretionary effort.

This research design in this study has several limitations that also offer avenues for future research. A very short time dimension is considered in this experiment. It is likely that over time individuals will change their behavior with respect to monitoring; however some research suggest that the change could be decreased tolerance rather than acclimation to intense monitoring (Smith et al. 1992). Although this research design has

limitations, it has the opportunity to open up new lanes of research on the topics of monitoring, control, and performance.

CHAPTER 3

THE EFFECTS OF MONITORING ENVIRONMENT ON PROBLEM SOLVING ABILITY

INTRODUCTION

In a successful organization, employees and managers should be creative and mentally flexible. This is especially true in the public accounting profession, where change is constant and the demand for services is high. Bonner and Lewis (1990) described "problem solving adaptability," in the auditing context, as the ability to recognize relationships, interpret data, and reason analytically. Similarly, Baril et al. (1998) claim that the ability to recognize that there are a variety of solutions to a particular problem is important to success in the accounting profession.[3] While researchers have tended to focus on individual characteristics correlated with problem solving (Gibbins and Jamal 1993), Libby and Luft (1993) warn that research which fails to consider the environment will miss important determinants of performance, since environmental factors affect motivation, knowledge, and ability. One increasingly important, but often overlooked, environmental variable is worker autonomy.

Some researchers have suggested that strict environmental controls can have detrimental effects on creative thinking (Hennesey and Amabile 1998; Elsbach and Hargadon 2006), which is critical for finding the optimal solution to complex problems.[4] Individuals are likely to be most creative when they experience high levels of intrinsic motivation (Amabile 1996), since such motivation increases their tendency to be curious,

[3] Many researchers have suggested that the ability to recognize and/or check for multiple solutions may be critical to auditor performance (e.g., Hieman 1990; Bierstaker at al. 1999).

[4] Luchins (1942, 37) was one of the first researchers to suggest that authoritarian control may increase problem solving fixation.

cognitively flexible, risk taking, and persistent in the face of barriers (Utman 1997; Zhou and Shalley 2003). Similarly, the theory of social facilitation states that individuals have various social, physiological, behavioral, and cognitive reactions to being monitored, watched, or judged (Zajonc 1965; Aiello and Douthitt 2001), which undermine performance on complex tasks, but positively affect performance on simple tasks. The purpose of the current research is to extend the managerial literature by investigating the effects of the monitoring environment on key aspects of individuals' problem solving ability.

In this research design, I use an adaptation of Luchins's (1942) water-jar task[5] to examine pattern establishment, problem solving rigidity (pattern breaking), and problem solving creativity, under three different types of worker monitoring. One hundred fourteen participants were assigned to one of three monitoring treatments, trust (no monitoring), human monitoring, and electronic monitoring. Once the treatment was induced, the participants were given the water-jar problems to test the three different aspects of their problem solving ability. Using computer illustration, the water-jar task gives the participant three water jars (Jar A, Jar B, and Jar C) of different sizes and asks them to fill one of those jars to a specific volume, a volume not directly available by filling only one of the jars (Appendix F shows the water-jar task user interface and the bottom of Appendix G shows the participant's answer form). Participants are then instructed to use the simplest method possible to solve each problem. They are allowed

[5] Fessler (2003) is one of the few research experiments in the recent managerial accounting literature to use the water-jar tasks. His research focused on task attractiveness, compensation scheme, and performance. In earlier accounting research, Stedry (1960) used the water-jar task to look at goal difficulty and goal acceptance.

two minutes to solve each problem by writing their answer on the answer form, but can move on to the next problem if finished before the time limit.

The first several problems (Problems 1- 6) all have the B-A-C-C solution as the correct answer. These problems test the participant's ability to establish and rely on a pattern. The next three problems (Problems 7-9) offer a simpler solution (either *A-C or A+C*) as well as the B-A-C-C solution. These dual answer problems are known as "critical problems."[6] Normally, most individuals who have found and used the pattern will become blind to the simpler solutions available to solve the critical problems, and will continue to use the pattern answer. Luchins (1942) labeled the solution blindness "Einstellung" effect.[7] These problems (Problems 7-9) test the participant's susceptibility to Einstellung blindness. The final three problems (Problems 10-12) all have different solutions, which increase in complexity, and test the participant's ability to solve complex problems. Appendix G shows a summary of the problem types the participants face as they progress through the water-jar task, and the answer sheet provided.

Based on past literature, I hypothesize that participants in the monitored treatments will 1) have more difficulty finding and using the pattern, 2) more often fail to recognize when the pattern is no longer efficient (Einstellung blindness), and, 3) solve fewer complex problems than individuals in the trust treatment. I find evidence to support the first two hypotheses related to pattern establishment and pattern breaking, and less clear, but interesting, evidence concerning the third hypothesis related to complex problem solving. With respect to complex problem solving, an interesting interaction

[6] For examples and more explanations of the terminology associated with the water-jar task, see Schultz et al. (1997).

[7] Einstellung means "setting" or "to set" in German.

between an individual's self-reported base ability and the monitoring environment was found. Individuals who reported being good in the related ability underperformed on complex problem solving in a trusting environment and outperformed in the human and electronic monitored environments.

THEORY AND BACKGROUND

Problem solving rigidity and Einstellung blindness are not just laboratory phenomena but are common cognitive biases.[8] Research suggests that even people who are professionals in their domain can miss simple solutions to "critical problems."[9] For example, Bilalić et al. (2008) showed that master chess players can become fixated to complex chess strategies after performing the same moves several times, and miss simpler effective moves. Similarly, evolutionary scientist Stephen Jay Gould (1996) discussed, in his controversial book *The Mismeasure of Man*, how scientists can be so strongly influenced by a theory they already hold, and have experience with, that they do not interpret new data objectively. In an accounting domain example, Marchant et al. (1991) found that expert tax preparers were slower to adapt to new tax laws because they failed to consider new relevant information.

Problem solving and creativity have been examined in conjunction with a variety of environmental and personal factors. In examples of environmental factors, researchers have shown that individuals can become rigid in their problem solving abilities when in stressful situations (Schultz and Searleman 1998; Cowen 1952) or faced with aggression (Carnevale and Probst 1998). Conversely, individuals who receive positive affect (Isen et

[8] Cognitive Bias is defined as a pattern of deviation in judgment that occurs in particular situations, which may sometimes lead to perceptual distortion, inaccurate judgment, illogical interpretation, or irrational behavior (Ariely 2009).

[9] "Critical problems" in this context is defined as problems with multiple solutions.

al. 1987) or who work on complex tasks with autonomy have been found to develop more creative ideas (Hatcher et al. 1999; Tierney and Farmer 2004). In examples of personality factors, researchers have shown that the "five factor" personality dimension of "openness" is associated with creativity (Fiest 1998) while Erikson (2012) found that "thinking" types of personalities appear to be resistant to problem solving rigidness and fixation.[10]

Financial and managerial accounting researchers have been interested in "fixation" for some time. A popular line of research is concerned with why some individuals appropriately update their decisions in response to changes in accounting methods and some individuals do not (Wilner and Birnberg 1986). The failure to adapt decisions to a change in accounting method is referred to as *accounting fixation.* Accounting fixation is indicated by the inability of users of accounting information to look behind the labels attached to accounting numbers (such as "cost" or "income") to adjust for changes that have occurred in the accounting techniques or methods used to determine that number (Bloom et al. 1984). Dearman and Shields (2005) argue that the ability to adapt one's decision process to a change in accounting method will be a function of one's task-relevant knowledge, problem solving ability, and intrinsic motivation. However, Wiley (1998) suggests that experts may be more prone to fixation because their domain-specific knowledge hinders their search for new information.

Given the role of auditors and public accountants in dealing with fraud and misreporting, understanding cognitive biases that hinder hypothesis generation and problem solving adaptability is critical. For instance, Bierstaker and Wright (2001)

[10] See Shalley et al. (2004) for review of contextual factors, personal factors, and interactions that affect creative thinking skills.

showed that practical problem solving ability was significantly correlated with performance on internal control evaluation tasks. Bonner and Lewis (1990) assert that problem solving ability, which they describe as the ability to recognize relationships, interpret data, and reason analytically, is one of the main determinants of auditor expertise. Further, SAS 56 calls for auditors to identify unexpected patterns in financial data and to hypothesize about the likely cause. However, auditors may become fixated on past patterns, and hypotheses, causing them to fail to incorporate all the new information they have at their disposal (Mock and Wright 1993; Wright and Bedard 2000). The failure to incorporate new data may lead to lower quality audits and, ultimately, more financial misstatements and financial fraud.

Problem solving rigidity works against the auditor's ability to solve unique problems and to reason analytically. In related examples, Mock and Wright (1993), looking at 74 random audits, found a weak relationship between client risks and audit programs. They also found that audit programs changed little over time with many tests done across a broad array of engagements. This is troubling since auditors are supposed to avoid becoming predictable, lest management and employees easily avoid the auditor's tests.

Asare and Wright (2003), in a laboratory experiment, examine the linkage among the initial hypothesis set, the information search, and decision performance in analytical procedures. They find that auditors who inherit, or generate, an incorrect hypothesis set, but still receive balanced evidence, do not perform well because they are unwilling or unable to generate additional hypotheses during the investigation phase, and they have difficulties integrating evidence that does not fit with their current hypothesis set. While

Asare and Wright (2003) acknowledge that auditors become fixated on certain hypotheses, they do not address the causes of fixation. Research into the environmental and personal factors which cause auditors or managers to become fixated on hypotheses and solutions may be important to the advancement of techniques and practices.

In this research, I investigate the effects of monitoring environment on different aspects of problem solving ability and problem solving fixation. If monitoring affects problem solving, the effects are likely attributable to social facilitation. The theory of social facilitation states that individuals have various social, physiological, behavioral, and cognitive reactions to being monitored, watched, or judged (Zajonc 1965; Aiello and Douthitt 2001). One common finding in the social facilitation literature is that monitoring negatively affects performance on complex tasks, and positively affects performance on simple tasks.[11] Since problem solving tasks are inherently complex,[12] there should be a negative relationship between monitoring intensity and most aspects of problem solving capabilities.

HYPOTHESIS DEVELOPMENT

The water-jar task can be broken down into three parts: the first involves pattern recognition and use, the second Einstellung blindness, and the third complex problem solving ability. Accordingly, I test for the effect of the monitoring environment on each part of the task, as shown in Appendix G. In the first part, the participant will establish a pattern, known in the problem solving literature as a mental set. When a mental set is established, the participant will tend to stop looking for new solutions and rely on the known pattern. Research suggests that high stress and low intrinsic motivation may

[11] The causes of social facilitation are thought to be primordial (Blascovich et al. 1999).

[12] Bonner et al. (2000, 25) describe problem-solving tasks as being complex.

activate individuals' need for closure, increasing the use of heuristics (Kruglanski and Fruend 1983) and leading individuals to lessen their search for new information (Klein and Webster 2000). Establishing a mental set usually leads to an increase in structure and a decrease in uncertainty (Shultz and Searleman 1997). If one is not enjoying the problem solving process, they will likely engage in the problem solving mental process for a shorter period of time, and will likely not stay open to the possibility of new solutions. With simple tasks, higher levels of monitoring should increase pattern recognition, pattern reliance, and the use of shortcuts.

However, the water-jar task is complex,[13] and as such, may be negatively affected by monitoring. In the presence of other people, by the phenomena known as social facilitation,[14] people tend to improve performance on simple tasks but their performance is impaired on complex tasks (Zajonc 1965, 1980). Aiello and Shao (1993) extended the social facilitation theory to include electronic monitoring by showing that when a task is the least bit complex (i.e., requires some thought) electronic monitoring lowers performance. In the water-jar task, being able to recognize and use the pattern solution requires complex thought. It is likely that individuals who are monitored at higher levels will not recognize and use the pattern efficiently, causing them to miss the pattern solution problems more often than individuals who are not monitored. Therefore I make the following substantive hypothesis:

H1: Individuals who are monitored at a higher level will fail to recognize and use a pattern more often than those who are not monitored at a higher level.

[13] Fessler (2003, 165) establishes that the water-jar task is considered a complex task.

[14] Bond and Titus (1983) provide a meta-analysis of 241 studies related to social facilitation.

In the second part of the water-jar task, after working six problems where the pattern solution was the best answer, the participant will be tested for Einstellung blindness with the next three problems. These problems offer the pattern solution from the earlier problems, but offer a simpler solution as well. The simpler solution is actually the correct solution, as the instructions are to give the simplest solution possible as their answer. As noted above, researchers have found that time constraints and stress lead to an increase in Einstellung blindness (Luchins 1942; Schultz and Searleman 1998). Thus, individuals who feel monitored or controlled will suffer from Einstellung blindness at a higher rate than those who do not feel monitored. This leads to the second hypothesis:

H2: Individuals who are monitored at a higher level will have Einstellung blindness more than those who are not monitored at a higher level.

In the third part of the water-jar task, the final three problems, the participants will have their analytical and complex problem solving skills tested by attempting to solve problems which have only complex solutions, and which increase in difficulty from one problem to the next. Creativity, which is essential for complex problem solving, is defined as the ability to transcend traditional ideas, rules, patterns, relationships, or the like, and to create meaningful new ideas, forms, methods, and interpretations.[15] Creativity in this context will be even more difficult because research has shown that once a mental set has been established in past problems, one will likely continuously attempt to apply the rule until they have a moment of insight (see Knoblich et al. 2008) and realize that that past rules are ineffective (Fantino et. al 2003).

Researchers have noted that when individuals use an established rule to solve problems, they have difficulty solving problems when the established rule no longer

[15] This definition is from www.dictionary.com.

works as a viable solution (Smith and Blankenship 1991). For example, Schultz and Searlemen (1998) found that, under stress, participants took an average of 43 seconds to work the pattern solution problems once they had established a pattern, but it took them an average of 64 seconds to solve simpler problems that did not have the pattern solution, after they were accustomed to the pattern solution. Additionally, 26% of their participants could not solve the simpler problems at all, in the time allotted. Given the time constraints in the current study (120 seconds) and the fact that all the previous problems can be answered using one (if the participant has Einstellung blindness) or two (if the participant does not have Einstellung blindness) simple rules, the added stress of being monitored should decrease the participant's ability to solve complex problem. This leads to the third hypothesis.

H3: Individuals who are monitored at a higher level will solve fewer complex problems than those who are not monitored at a higher level.

RESEARCH DESIGN

This experiment was carried out in a computer lab at large public university. Twelve sessions were held, each with either 9 or 10 participants. A diverse group of 114 adult volunteers participated, recruited through the university email newsletter, flyers, and word of mouth. Participants self-registered online and chose the session date and time they preferred. Upon arrival, they were given a short demographic questionnaire, as shown in Appendix A.

The sessions were pre-assigned to one of three monitoring treatments: electronic monitoring as the most intense treatment,[16] human monitoring as a mid-level treatment, and trust (no monitoring) as a low-level treatment. When the participants arrived they were provided a consent form and demographic questionnaire. Table 3.1 shows the demographics of the participants. Next, their attention was directed to the white board where the task schedule and compensation plan were explained. With respect to the task schedule, the participants worked on a data-correction task, for a flat wage, for approximately 27 minutes, before working on the water-jar task. The prior task was used to acclimate the participants to their monitoring treatment and the environment in general, and to collect data for a separate study.

[16] While electronic monitoring and traditional human monitoring have the same fundamental purpose, past research suggest that the pervasive, continuous nature of electronic monitoring often elicits stronger reactions from the worker (Aiello and Kolb 1995; Lund 1992; Stanton 2000).

Table 3.1
Demographics by Treatment Group

Gender	Trust Monitored	Human Monitored	Electronically Monitored	Total	Percent
Male	19	17	18	54	47%
Female	19	21	20	60	53%
				114	100%
Age					
18-24	23	26	18	67	59%
Over 24	15	12	20	47	41%
				114	100%
Student Nationality					
Domestic	23	24	21	68	60%
International	14	10	12	36	32%
Not a Student	1	4	5	10	9%
				114	100%
Business Student					
Business Student	14	11	11	36	32%
Non-Business Student	23	23	21	67	59%
Not a Student	1	4	6	11	10%
				114	100%
College Level					
Fresh/Soph	19	18	8	45	39%
Junior/Senior	13	8	12	33	29%
Graduate	5	8	10	23	20%
Not a Student	1	4	8	13	11%
				114	100%
Good at Mental Math					
Yes	23	29	20	72	63%
No	15	9	18	42	37%
				114	100%

Note: This table shows the key demographics for each treatment group. The Mental Math category shows the response to the statement, "I consider myself good with mental math and numbers."

Participants in the electronic monitoring treatment had a small webcam facing their workstation keyboard and papers when they arrived, as shown in Appendix B. Participants in this electronically monitored group received all the same task instructions as the other two groups, except they were told "you are being monitored with webcams so we can observe your work and make sure you follow the instructions as given."[17] Unbeknownst to the participants, the webcams were not activated. However, the experimenter and research assistant sat prominently at a corner workstation that the subjects believed to be the "monitoring station." The station was manned for the entire task period.

Participants in the human-monitored group received all the same task instructions as the other two groups, except they were told that "I will walk around the room so I can observe your work and make sure you follow the instructions as given." The researcher and the research assistant wandered around the room and passively observed the participants during the task time.

Participants in the trust treatment received all the same task instructions as the other groups, but were told "you will not be watched and we believe you will follow instructions as given." In this treatment, all research personnel left the room and returned when time was up.

When the water-jar task began, the students were given answer sheets and instructed to open the water-jar task slideshow on their computer. The researcher then introduced the participants to the task and walked them through the first two examples, which had the answers already entered on the answer sheet, as shown at the bottom of

[17] This wording used in the participant instructions is based on the wording used in the study by Enzle and Anderson (1993). Their study looked at the effects of controlling versus non-controlling electronic monitoring on intrinsic motivation.

Appendix G. The last two examples were worked with the participants, who had to enter the answers on the answer sheet. The researcher made sure that everyone entered answers for the example problems correctly before allowing the problem portion of the task to begin. The four example problems ensured that each participant had at least a basic understanding of how to complete the task before being allowed to proceed to the actual problems.

The actual experimental task consisted of twelve problems. Participants had a two minute time limit on each problem[18] and were instructed to move on if they ran out of time, or as soon as they solved the problem. They were told that they would be compensated based on the number of correct solutions they gave, and that if more than one solution to a problem was found then the simplest solution, the one with the least moves, was the correct one. If they finished the task quickly they were to sit quietly until time was up.

RESULTS

Pattern Recognition

Hypothesis 1 stated that individuals who are monitored at a higher level will more often fail to recognize and use a pattern than those who are not monitored at a higher level. In this experiment, one was considered to have recognized and used the pattern if they correctly answered at least 5 of the 6 pattern problems given in the task (the first six problems had the B-A-C-C[19] solution). With respect to monitoring levels, as discussed

[18] The participants had two minutes to solve each problem. At the bottom of each problem was a timer bar that filled up, to mark the time. When the bar became full a bell sounded and the words "click to the next slide" appeared across the screen.

[19] Read as "B minus A minus C minus C."

above, electronic monitoring was the most intense treatment, human monitoring was a mid-level treatment, and trust (no monitoring) was a low-level treatment.

Out of the 114 participants, 79 recognized and used the pattern solution correctly, while 35 did not. Table 3.2, Panel A, shows summarizes the number of participants who recognized the pattern, by treatment. Seven people in the trust treatment failed to recognize the pattern, 12 in the human monitoring treatment, and 16 in the electronic monitoring treatment. A Chi-square analysis of the table (bottom of Panel A) shows that pattern recognition was not independent of monitoring treatment (p= .081). In addition the effect seems to occur exactly as hypothesized, with pattern recognition negatively related to monitoring intensity. Using Somers' d, a directional test of association of one ordinal/ranked measure (treatment) and one nominal measure (pattern recognition), shows the relationship is significant (p=.02; seen at the bottom of Panel A).

Self-perceived mental math ability is positively associated with pattern recognition (p= .01; analysis not shown) and the treatment groups were imbalanced with respect to mental math ability, as shown in Table 3.1. Of individuals in the electronic monitoring treatment, 47% indicated on the demographic questionnaire that they were not good at mental math, compared to 24% for the trust treatment, and 39% for the human monitored treatment. To consider mental math ability and monitoring treatment together, a logistic regression is used where the dependent variable is binary for pattern recognition, and monitoring treatment and mental math ability (binary) are the independent variables. Panel B shows that, after controlling for mental math ability, human monitoring (versus the benchmark of the trust group) decreases pattern

recognition (p= .096), and electronic monitoring decreases pattern recognition (p= .012). Therefore, Hypothesis 1 is supported.

Table 3.2
Pattern Recognition by Treatment Group

Panel A: Pattern Recognition by Treatment

	Recognize Pattern		
Treatment	**No**	**Yes**	**Total**
Trust	7	31	38
Human Monitoring	12	26	38
Electronic Monitoring	16	22	38
Total	35	79	114

Pearson Chi² = 5.03 **Pr=** *0.081*
Somers' d = -0.158 **Pr=** ***0.020***

Panel B: Logistic Regression of Treatment and Mental Math on Pattern Recognition

		Observations	114
Log likelhihood	-64.49	**LR chi²**	5.440
Pseudo R2	0.083	**Prob > chi²**	0.001

Variable	**Coef.**	**Std. Error**	**Z**	**P>|z|**
Mental Math	1.11	0.44	2.51	**0.012**
Electronic Monitoring	-1.15	0.55	-2.10	**0.036**
Human Monitored	-0.95	0.57	-1.66	*0.096*
Constant	0.91	0.47	1.93	*0.053*

Note: Panel A of this table shows the number of participants who recognized the pattern solution of the water-jar task, by treatment group. A participant is considered to have recognized the pattern if they got at least six of the seven pattern problems correct. Panel B shows a logistic regression of the data in Panel A but also controls for the participant's self-reported mental math abiltiy (binary), as reported on the demographics questionnaire. P-values < .05 are **bolded**; those < .10 are *italicized.*

Einstellung Blindness

Hypothesis 2 stated that individuals who are monitored at a higher level will have Einstellung blindness more than individuals who are not monitored at a higher level. In this experiment, an individual is considered to have Einstellung blindness if they answer Problem 7, the first dual answer problem, with pattern answer (B-A-C-C) instead of the simpler solution (A-C). The simpler solution was the correct solution, as the participants were instructed to use the simplest answer available.

Of the 114 participants, 51 failed to see the simpler solution and instead used the pattern solution (Einstellung blindness), 55 properly used the simpler solution, and eight used neither the simpler solution nor the pattern solution (missed the problem completely). If Einstellung blindness is the cause for roughly half of the participants continuing to use the pattern solution, then there should be a positive correlation between pattern recognition/use and Einstellung blindness, since one could not be blinded by a pattern they were not relying upon. Panel A of Table 3.3 confirms this proposition. Of the 51 participants who were blind to the simpler solution, 45 recognized and used the pattern solution correctly and 6 did not. Of the 55 participants who were not blind to the simpler solution 33 recognized and used the pattern solution correctly and 22 did not. A Chi-square analysis (bottom of Panel A) shows that the effect of pattern recognition on Einstellung blindness is significant ($p = .001$).

Panel B of Table 3.3 shows the number of Einstellung blindness participants by treatment, isolated to those who recognized and used the pattern solution. The table shows that the participants in the trust treatment were less susceptible to Einstellung blindness than those in the other two treatments. In all, 51% (17/33) of those who recognized the pattern but did not suffer Einstellung blindness were in the trust treatment,

while 31% (14/45) of those who recognized the pattern but did suffer Einstellung blindness were in the trust treatment. A Chi-square analysis of the table (bottom of Panel B) shows that Einstellung blindness was not independent of monitoring treatment (p= .022). However, the effect does not seem to occur exactly as hypothesized (with pattern recognition negatively related to monitoring intensity). It seems that the Einstellung blindness is more likely in the human monitoring than in the electronic monitoring. Using Somers' d, as a directional measure of association of one ordinal/ranked measure (treatment) and one nominal measure (Einstellung blindness), shows the relationship is not significant (p= .478) when monitoring intensity is ranked (bottom of Panel B).

Panel C of Table 3.3 shows the results for a logistic regression examining the effects of the monitoring treatment on Einstellung blindness (binary) while controlling for the effects pattern recognition (binary) and self-reported mental math ability (binary). The results show that mental math ability is negatively related to Einstellung blindness (p= .10), while pattern recognition is positively related to Einstellung blindness (p< .001). The human monitoring treatment is positively associated with Einstellung blindness (p= .005) while the there is little difference in the other two groups.[20]

[20] The regression in Panel C of Table 3.3 yields similar results, with respect to the monitoring treatments, when using only the 78 participants who recognized the pattern, and omitting pattern recognition as an independent variable.

Overall, the evidence supports Hypothesis 2, which states that individuals who are monitored at a higher level will have Einstellung blindness more than those who are not monitored at a higher level. While questions remain as to why participants in electronic monitoring treatment were less susceptible to Einstellung blindness than the human monitoring treatment, there is little doubt that monitoring affected this area of problem solving.

Table 3.3
Einstellung Blindness by Pattern Recognizers and Treatment Group

Panel A: Blindness by Pattern Recognition

	Einstellung Blindness		
Pattern Recognition	**No**	**Yes**	**Total**
No	22	6	28
Yes	33	45	78
Total	55	51	106

Pearson Chi² = 10.85 **Pr= 0.001**
Fisher's Exact **Pr= 0.002**

Panel B: Einstellung Blindness by Treatment for Pattern Recognizers

	Einstellung Blindness		
Treatment	**No**	**Yes**	**Total**
Trust	17	14	31
Human Monitoring	5	20	25
Electronic Monitoring	11	11	22
Total	33	45	78

Pearson Chi² = 7.62 **Pr= 0.022**
Somers' d = 0.69 **Pr=** 0.478

Panel C: Logistic Regression of Treatment, Mental Math, and Pattern Recognition on Einstellung Blindness

		Observations	106
Log likelhihood	-62.14	**LR chi²**	22.510
Pseudo R2	0.153	**Prob > chi²**	0.000

Variable	**Coef.**	**Std. Error**	**Z**	**P>|z|**
Mental Math	-0.81	0.49	-1.64	*0.100*
Pattern Recognition	2.09	0.59	3.55	**0.000**
Electronic Monitoring	0.19	0.53	0.35	0.724
Human Monitored	1.61	0.57	2.81	**0.005**
Constant	-1.71	0.65	-0.27	**0.008**

Note: Panel A and B of this table show the number of participants who were susceptable to Einstellung blindness, by pattern recognition and treatment, respectively. A participant is considered to have Einstellung blindness if they gave the pattern answer to Problem 7 instead of the simpler answer. Panel C shows a logistic regression of the data in Panel A and B, but also controls for the participant's self-reported mental math abiltiy (binary), as reported on the demographics questionnaire. P-values < .05 are **bolded**; those < .10 are *italicized.*

Complex Problem Solving

Hypothesis 3 states that individuals who are monitored at a higher level will solve fewer complex problems than those who are not monitored at a higher level. In this experiment, the last three problems had increasingly complex answers. The answers to Problem 10, 11, and 12 were B-A+C, B-C-C+A, and A+C+C+C+C, respectively.[21] Few participants answered all three of these problems correctly, given the time constraints (120 seconds) and the fact that all the previous problems could be answered using one or two simple rules.

Of the 114 participants, 9 (7.8%) did not answer any of the three correctly, 47 (41.2%) answered 1 of the three correctly, 43 (37.7%) answered two of the three correctly, 13 (11.4%) answered all three correctly, and 2 (1.7%) participants did not follow directions (spent more than 120 seconds on one or more of these three problems). However, Panel A of Table 3.4 shows little difference in the average number of complex problems solved across groups. Participants solved 1.49 complex problems on average in the trust treatment, 1.53 in the in the human monitoring treatment, and 1.59 in the electronic monitoring treatment. A linear regression of monitoring treatment on the number of complex problems solved, as shown in Panel B of Table 3.4, confirms that, even after controlling for mental math ability, monitoring treatment does not seem to have a strong effect on complex problem solving skills. Therefore Hypothesis 3, as stated, is not supported.

[21] Given that the order of operations does not differ for addition and subtraction, different versions of these answers were considered as if they were given. For example, B-C-C+A is equal to B+A-C-C.

Table 3.4
Complex Problem Solving by Treatment Group

Panel A: Summary Statistics for Complex Problem Solving

Treatment	Obs.	Mean	Std. Dev	Min	Max
Trust	37	1.49	0.768	0	3
Human Monitoring	38	1.53	0.762	0	3
Electronic Monitoring	37	1.59	0.900	0	3
Total	112	1.54	0.805	0	3

Panel B: Regression Results for the Effects of Monitoring Treatment and Mental Math Ability on Complex Problem Solving

Number of Observations	112	**R- squared**	0.01
Root MSE	0.810	**Adj R-squared**	-0.01

Source	Coef	Std Err	t	P>|t|
Mental Math	0.171	0.163	1.05	0.295
Human Monitoring	0.016	0.189	0.08	0.934
Electronic Monitoring	0.127	0.189	0.67	0.505
Constant	1.380	0.167	8.25	**0.000**

Panel C: Regression Results for the Interaction of the Monitoring Treatment and Mental Math Ability on Complex Problem Solving

Number of Observations	112	**R- squared**	0.10
Root MSE	0.783	**Adj R-squared**	0.05

Source	Coef	Std Err	t	P>|t|
Mental Math	-0.481	0.265	-1.81	*0.073*
Human Monitoring	-0.563	0.335	-1.68	*0.095*
Electronic Monitoring	-0.508	0.279	-1.82	*0.072*
Mental Math*Human Monitoring	0.880	0.400	2.20	**0.030**
Mental Math*Electronic Monitoring	1.100	0.370	2.97	**0.004**
Constant	1.790	0.209	8.53	**0.000**

Note: Panel A shows the summary statistics for the number of complex problems solved by each treatment group. Panel B shows the results for a regression of self-reported mental math abiltiy (binary), as reported on the demographics questionnaire, and monitoring treatment on complex problem solving. Panel C Shows the interaction of mental math ability and monitoring on complex problem solving. P-values < .05 are **bolded**; those < .10 are *italicized* .

Past research suggests, however, that there may be an interaction of perceived ability and different types of monitoring regimes on complex problem solving ability. For instance, Davidson and Henderson (2000) found that the visual presence of electronic monitoring resulted in an easy task being performed with greater proficiency and a difficult task being performed with less proficiency. In their study, when participants tried to solve an easy task, the presence of electronic monitoring resulted in their mood state becoming significantly more positive; but when solving a difficult task, electronic monitoring caused a more negative mood state. Therefore, for robustness, an interaction effect between monitoring treatment and mental math ability on complex problem solving is considered.

Panel C of Table 3.4 shows that when interactions between self-perceived mental math ability and monitoring treatments are considered in a regression, they are highly relevant. The positive interaction terms indicate that, for persons higher on self-perceived mental math ability, monitoring increases their performance. Hence, the effect of monitoring on complex problem solving depends on the person's perceived mental math ability.

Figure 3.1 shows the predicted marginal means for the interaction. Interestingly, the results suggest that individuals who thought they were good at mental math and were monitored (human and electronic) performed very well, while those who thought they were good at mental math and were not monitored (trust) performed poorly. Conversely, individuals who were less confident in their base ability and in the trust treatment (left alone) performed very well, but such individuals who were monitored (human and electronic) performed poorly. One possible explanation for this finding is that those who

are confident like to be watched, possibly lowering stress or making the task seem more attractive, while those who are not confident do not like to be watched, as it may raise stress or make the task seems less attractive. More research should be done to determine the relationship of ability, monitoring, and performance.

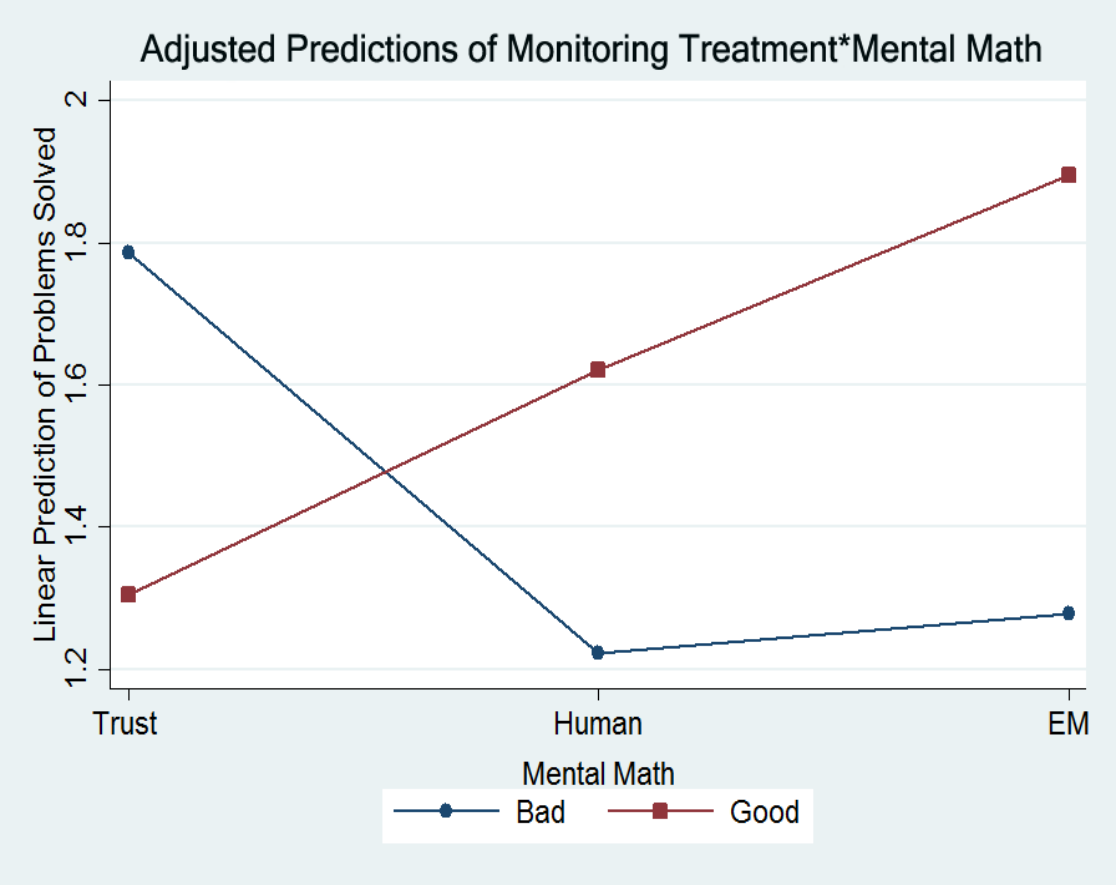


Figure 3.1: Adjusted Marginal Means for the Interaction of Monitoring Treatment and Self-reported Mental Math Ability on Complex Problem Solving

CONCLUSION

Creative problem solving is increasingly important in many business domains. While much is known about personal characteristics that are important for creative problem solving, less is known about how the environment affects an individual's problem solving abilities. As such, organizations may choose their internal control systems without fully understanding the effects the controls will have on their employees' ability to solve problems. I add to the current literature by investigating the effects of three different monitoring regimes (trust monitoring, human monitoring, and electronic monitoring) on three different areas of problem solving (pattern recognition and use, Einstellung blindness [pattern breaking], and complex problem solving skills).

The results indicate that the monitoring environment does influence individuals' problem solving. For the most part, problem solving performance is negatively related to monitoring intensity. However, when solving complex problems, if an individual is confident in his or her base skills, then the negative effects of monitoring intensity may be mitigated or even reversed. These findings are consistent with the theory of social facilitation, which states that individuals have various social, physiological, behavioral, and cognitive reactions to being monitored, watched, or judged (Aiello and Douthitt 2001). One common finding in the social facilitation literature is that monitoring undermines performance on complex tasks, and improves performance on simple tasks. More research should be done to determine how the awareness of social facilitation can be incorporated into management theory, and ultimately be used to help to improve job training, job design, and job performance.

With respect to accounting and auditing, this paper discusses an often overlooked phenomenon that may contribute to lower audit quality, increased misstatements, and

fraud—Einstellung blindness. Einstellung blindness is a stealth cognitive bias related to problem solving rigidity and functional fixation. In short, Einstellung blindness occurs when one has been exposed to the same type of problem and solution so many times that they become cognitively blinded to any other solution. While not widely studied in recent business literature, the impact of Einstellung on various business- and accounting-related job duties (e.g., auditing, tax compliance, and managerial decision making) could be substantial. In this paper I show that participants who are monitored at a higher level have Einstellung blindness more often than those who are not monitored at a higher level.

Individuals in the auditing profession may be especially susceptible to Einstellung blindness for at least three reasons. First, auditors often repeat similar tests which yield similar results. Second, auditors are often under time pressure. Third, auditors usually have their work monitored, or are subject to monitoring during their audit tasks. More research should be conducted to examine if, and how, Einstellung blindness affects audit quality. A similar line of thought can be extended to other areas of the accounting and managerial domains.

Lastly, while a wealth of interesting research exists on the effects of financial incentives on performance and motivation in the managerial accounting literature (Bonner et al. 2000), there has been considerably less research on the effects of monitoring and control on motivation and performance. Since monitoring and incentives are thought to be the two main sources of organization control (according to traditional agency theory), more work should be done to develop similar knowledge on each topic. In this aspect, this research answers Christ et al.'s (2012) call to further develop our understanding of the potential consequences of formal controls on the agent's behavior.

CHAPTER 4

DOES MONITORING AFFECT THE AGENT'S PREFERENCE FOR HONESTY?

INTRODUCTION

It is firmly established in the business literature that monitoring increases effort and deters dishonest behavior within a firm (e.g., Hölmstrom 1979; Jensen and Meckling 1976). This conclusion is logical, and rational, since any self-interested agent should work hard and be honest to avoid the possibility of sanctions if caught shirking or being dishonest. Despite the importance of monitoring in the firm, there is little research in the accounting and managerial literature addressing the effects of monitoring and control on the individual psyche.[22] For instance, how monitoring affects the agent's attitude toward dishonesty and misreporting is largely an unanswered empirical question. This is an important issue because past research has shown that attitude is highly correlated with intent and future behavior (Ajzen and Fishbein 2005). Since attitude/rationalization is considered one of the three sides of the fraud triangle (PCAOB 2005; Cressy 1973), understanding the relationships among monitoring, attitudes, and (dis)honesty is vital to the design of internal controls, financial regulation, and the prevention of fraudulent behavior.

This study proposes that monitoring negatively affects the agent's attitude towards honest reporting by "crowding out" the agent's intrinsic motivation to be honest and enabling the rationalization of deviant behavior. This theory is tested by experimentally investigating whether the type of monitoring affects an individual's

[22] For a broad review of honesty in managerial research see Salterio and Webb (2006).

behavioral honesty. In this experimental design, participants were assigned to one of three monitoring treatment groups: a trust treatment,[23] a human monitored treatment, [24] and an electronically monitored treatment.[25] Once the treatment was induced, the participants performed a simple mental math task where a monetary reward was given based upon task performance. Half the participants in each treatment group self-reported their results, while the other half had their results verified by the researcher. Dishonesty was operationalized by examining the difference in means between the "self-score regime" and the "verify regime" of each treatment group (see Ariely et al. 2009 for a similar research design). As hypothesized, the results of the experiment show that there was more dishonesty in the human monitored treatment and the electronically monitored treatment than the trust treatment. Interestingly, less dishonesty was detected in the electronic monitoring treatment than the human monitored treatment.

Psychology research suggests that individuals can be either internally or externally motivated to perform a task or carry out a behavior. When an individual already is intrinsically motivated to perform a behavior, controlling or incentivizing that behavior may externalize the motivation (Ryan and Deci 2000; Frey and Oberholser-Gee 1997). Externalizing intrinsic motivation can have negative consequences such that when the external control mechanism is removed or weakened the incentive to perform the behavior is diminished from its original state (Deci et al. 1999).

---

[23] Rousseau et al. (1998, 395) defined trust as "a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behavior of another."

[24] Human monitoring is sometimes referred to as "traditional monitoring" in the academic literature (e.g., Stanton 2000).

[25] While electronic monitoring and traditional human monitoring have the same fundamental purpose, past research suggest that the pervasive, continuous nature of electronic monitoring often elicits stronger reactions from the worker (Aiello and Kolb 1995; Lund 1992; Stanton 2000).

Researchers across multiple disciplines have found that most individuals are intrinsically motivated to be honest, and behave as if there is a "cost of lying" that must be covered before a lie is told (Lundquist et al. 2007; Gibson et al. 2012; Gneezy 2005). Although individuals behave as if there is a cost of lying, for most people that cost is not high; most individuals will lie for a small amount of gain (Gneezy 2005; Baiman and Lewis 1989). This suggests that there is a trade-off between being honest (internal gratification) and receiving a payoff by being dishonest. Empirical research suggests that personal characteristics and situational circumstances determine the point at which a lie becomes acceptable for each person.[26] Ariely et al. (2009) posit, in their theory of Self-Concept Maintenance, that individuals are only honest enough (partially honest) to convince themselves of their own integrity. They state that "a little bit of dishonesty gives a taste of profit without spoiling a positive self-view" (p. 3).This finding is consistent with several experiments which show that individuals are more likely to be a little dishonest than completely honest or completely dishonest. These two streams of research, together, suggest that individuals will lie for a small amount of gain, but will limit the impact of, or gain from, their lie to a certain threshold so that the lie does not alter their self-image. If monitoring negatively alters one's attitude toward honesty, facilitating rationalization of fraudulent behavior, then it is likely that it will also cause individuals to lower their threshold for dishonesty ("cost of lying"), and to the extent that rationalization allows one to be dishonest and still maintain their positive self-image, monitoring will also increase ones capacity for ill-gotten gains.

[26] Several researchers have looked at the causes of deviant behaviors such as lying and misreporting. Personal characteristics such as Machiavellianism (Fulmer et al. 2009; Murphy 2012) and self-control (Ariely et al. 2009) along with situational characteristics such as the business climate (Crutchley et al. 2007) and controls (Tayler and Bloomfield 2011) have been examined recently in the business literature.

By examining the effects of monitoring on behavior, this research answers Christ et al.'s (2012) call to further develop our understanding of the potential consequences of formal controls. Also, by positing that monitoring affects the participant's attitude toward misreporting, leading to rationalization, I heed the call of Hogan et al. (2008) to design studies in which multiple elements of the fraud triangle are examined simultaneously. As discussed above, the evidence suggests there may be a natural tension between the effects of control mechanisms and the externalization of intrinsic motivation. Thus, an attempt to reduce one side of the triangle (opportunity) through monitoring may weaken another side of the triangle (rationalization). The understanding of the relationship between these two opposing forces is important to the design of effective regulation and internal controls.

This study continues an interesting line of research in the accounting literature that is concerned with how the business environment can influence an individual's propensity to commit fraudulent or deviant behavior in accounting and managerial related domains. With regard to the "fraud triangle," researchers are interested in the rationalization and attitudes related to dishonest behavior, whether they are developed through the tone at the top (Rezaee 2005), contract design (Evans et al. 2001), the vertical and horizontal equity of compensation (Matuszewski 2010), personality traits (Murphy 2012), or other factors. The current research adds to the managerial accounting literature by investigating the possibility that monitoring, which is meant to prevent dishonest behavior, may actually promote dishonesty, under some circumstances, by making it easier for the agent to rationalize dishonesty.

The rest of this paper is organized as follows: Section II provides more theory and background, Section III develops the hypothesis, Section IV describes the research design, Section V provides the results, and Section VI gives the summary and conclusion.

THEORY AND BACKGROUND

For the most part, honesty in the accounting and finance literature is discussed in the framework of agency theory and/or fraud prevention. In both of these frameworks, monitoring is usually viewed in a positive light, where the only restraint on monitoring and control is the monetary limits of the principal (e.g., Zajac and Westphal 1994; Hansen 1997). However, some research suggests that that there are "hidden costs," and unanticipated effects, of monitoring and control. Some of these costs and effects derive from the fact that, given an acceptable option or alternative, people will choose not to be controlled. In other words, as a person's autonomy is removed, their internal motivation to cooperate with the authority is diminished (Spector 1986). However, the implicit costs of control are not well understood and are rarely considered in theoretical models. This paper addresses one dimension of these costs by looking at the effects of monitoring on the agent's behavioral honesty.

Honesty in Economics and Psychology

Honesty in the psychology literature is often contrasted with the view of honesty in the economics literature. The standard economic perspective of behavior is one of *homo economicus*, where the individual is a rational and selfish entity interested only in maximizing their own external payoffs. For *homo economicus*, the decision to be honest, or dishonest, depends only on the expected benefit versus the expected cost. This cost-benefit tradeoff means that decisions about honesty are like every other decision that

individuals face. For *homo economicus,* all else equal, an increase in reward will always increase a behavior, while an increase in punishment, or cost, will always decrease a behavior.

In contrast, the psychology literature holds that in addition to the external reward mechanisms, there also exist internal reward mechanisms and that these internal rewards influence individuals' decisions. The external and the internal reward mechanisms interact to determine if, and to what extent, an individual performs a behavior. From this interaction there is a non-linear relationship between honesty and the reward for being dishonest (see Ariely et al. 2008). However, because of differences in individual values, preferences, and cognition, the functional relationship between honesty and the reward for being dishonest seems to vary greatly between individuals and situations (Gibson et al. 2012).[27]

Gneezy (2005), in a simple game where one participant had the option to tell the truth or lie to another participant about the payoff from various options that they would split, found that the decision maker uses the "truth telling" outcome as a reference level when evaluating the benefits of lying. The monetary consequences of the lie are compared to this reference level. The decision maker is selfish in the sense of maximizing their own payoffs, but sensitive to the cost the lie imposes on the other side. Sensitivity diminishes with the size of payoffs. Moreover, since the perception of the counterpart's cost is subjective, when there are differences in wealth as in employee-employer relationship or a consumer-insurer relationship, the lower wealth decision maker is more likely to be dishonest.

[27] In the article titled "*In Search of Homo Economicus: Behavioral Experiments in 15 Small Scale Societies*," Henrich et al. (2001) test individuals from different types of societies to determine how much their decision making deviates from rational models.

Since some behaviors, such as an individual's concern for the counter-party, are not consistent with the characteristics of *homo economicus*, additional theories have been developed to help account for the discrepancies between economic-rationality and actual human behavior. For example, the theory of Self-Concept Maintenance (Ariely et al. 2008) posits that individuals are practically always in a win-lose situation where every decision is a trade-off between being honest and receiving an intrinsic reward or gaining from deception. However, instead of making a decision to be honest or dishonest, individuals usually look for a compromise. Individuals are often dishonest, but they limit their dishonest activity to a point where they do not have to change their own self-perception. The theory posits that the changing of one's self-perception is undesirable, or costly; but being partially honest offers the individual the "best of both worlds," gaining from dishonesty but still perceiving themself to be an honest and ethical person. The theory of Self-Concept Maintenance is pertinent to the study of "monitoring and honesty" because the type of monitoring may affect the internal threshold of dishonest behavior one can engage in and not have to update their self-identity.

Honesty and Agency Theory

Agency theory is the most comprehensive and widely accepted theory in managerial research and organizational design. Agency theory is useful in research, and practice, because it makes explicit predictions about how individuals are likely to behave under different contractual designs. To arrive at such predictions, agency theorists make assumptions about the people involved in the contracts, the entities offering and accepting the contracts, and the informational environment (Eisenhardt 1989). One of the main assumptions that agency theorist make about individuals is that they are rationally

self-interested, similar to *homo economicus*. Accordingly, a great deal of research has looked at relaxing this strict assumption (see Cuevas-Rodriguez et al. [2012] for a recent review). For example, researchers have found that the inclusion of trust (Beccerra and Gupta 1999), reciprocity (Kuang and Moser 2009), and social norms (Fehr and Falk 2002) into agency theory can dramatically alter the predicted outcomes of contracts.

Interestingly, experimental managerial accounting research has been a fruitful area for the study of behavioral agency theory models. Participative budget experiments, in particular, offer a unique setting where the information environment and/or the incentive structure of contracts, in the principal-agent relationship, can be manipulated and the effects of the manipulation on the agent's reporting and production decisions can be measured (Brown et al. 2009). This research is unique in the business literature because it allows researchers to empirically examine some determinants of honesty in an organizational setting. While more than two dozen published participative budget experiments in the managerial accounting literature over the past twenty years have addressed managerial reporting,[28] here, I review a few papers from a widely cited line of work that deals explicitly with honesty.

Evans et al. (2001) specifically examined how agents' preferences for honesty and wealth affect their reporting of private information. In their experiment the managers (participants) privately observe the cost of production and report it to the principal, who provides the amount requested. The agent keeps any surplus from over reporting and cannot be auditing or monitored. Interestingly, they found that, of the available surplus that the agent could have kept with impunity, the agents actually returned 47.6% through

[28] See Brown et al. (2009) for a review of participative budget experiments in the managerial accounting literature.

full or partial honesty. Evans et al. (2001) compare their results to the average of several dictator game experiments, where a participant simply decides how much of total sum to share with a person they have never met but has entrusted them with gains. In the dictator games the participants give back, on average, 18% of their gains. They attribute the difference, between budget experiment (47.6%) and dictator experiment (18%), to the fact that in the budget experiment the participant had to tell a lie to receive the surplus, in which case their preference for honesty, or partial honesty, affected their gain.[29]

Hannan et al. (2006) examine honesty in the participative budget setting under different levels of information asymmetry, while maintaining the trust setting from Evans et al. (2001). In their experiment the main variable was the precision of an information system (coarse or precise) that signaled the actual costs to the principal, although the principal has no power to deter dishonesty. They show that agents' reporting decisions are affected by how they trade off the psychological benefits of appearing honest against the economic benefits of misrepresentation. The precision of the information system affects the agent's trade-off by changing the ability of the principal to infer the agent's level of honesty. They find that honesty is lower under a precise information system than under a coarse information system because the incremental cost of appearing honest is higher with a precise system.

Rankin et al. (2008) extend the findings of Evans et al. (2001) and Hannan et al. (2006) by distinguishing more clearly whether agents' tendency to report private information more truthfully, despite an economic incentive to be dishonest, is due to honesty or to other non-pecuniary motivations such as fairness or reciprocity. They

[29] Fredrickson and Cloyd (1998) had similar findings and concluded, from agents' self-reported motivations, that personal integrity is the most important factor limiting slack in their experimental budgetary setting.

manipulate whether the agent's budget report does or does not require a factual assertion, noting that while fairness preferences could come into play in both conditions, honesty should come into play only when agents are required to make a factual assertion. They find more honest reporting when a factual assertion is required, indicating an incremental effect of honesty beyond other non-pecuniary preferences. In addition, Rankin et al. (2008) examine whether their finding holds when the principal rather than the agent has final budget authority. They find that the incremental effect of honesty is no longer significant when the principal has final budget authority. They also provide evidence suggesting that this may be because agents frame the situation as an ethical dilemma when the agent has final authority, but as a negotiation in which each party acts in his or her self-interest when the principal has final authority.

Overall, the evidence from the managerial accounting literature suggests that, all else equal, agents have a preference for partial honesty when there is a reward for lying. Agents will limit their dishonesty because they also have a preference for non-pecuniary benefits such as fairness, reciprocity, and honesty. The results of these experiments show the complexity of human decision making by suggesting that people "want their cake and they want to eat it too."

A common element of many of the participative budget experiments is the use of low monitoring to measure innate honesty. The study presented in this paper is unique because it attempts to measure how different monitoring environments affect honesty. I posit that monitoring makes it easier for agents to rationalize dishonesty when the opportunity arises, which may lead to an increase in dishonest behavior in environments where monitoring intensity is higher.

Honesty, Fraud, and Internal Controls

Fraud prevention and internal controls are concerned with preventing financial crimes, deterring misreporting, and safeguarding firm assets. Fraud involves intentional acts and is perpetrated by human beings using deception, trickery, and cunning (Ramamoorti 2008). Since fraud involves people's capacity to deceive, and be deceived, it is important to understand the psychological factors that might influence these types of behavior. Therefore, most work related to honesty and fraud prevention, or internal controls, is concerned with how and why individuals commit acts of fraud and deceit in the workplace or financial markets.

Despite increases in regulation and ethical training in the post Sarbanes-Oxley era**,** fraud and misreporting continue to be a pertinent threat to capital markets and internal controls (Hogan et al. 2008). Behavioral research, which exposes some shortcomings of theories based on economic rationality, suggests regulation and punishment may not affect decision making as much as previously thought. Further, ethical training may not be as effective if individuals delude themselves of their moral identity as the theory of Self-Concept Maintenance suggests. In light of the increases in reported fraud and financial crime, regulators have called for more research on the how to prevent or detect fraud (Hogan et al. 2008).

In their 2009 Global Economic Survey, PricewaterhouseCoopers described fraud and misreporting as pervasive, persistent, and pernicious.[30] Thirty percent of the 3,037 respondents reported dealing with either fraud or misreporting, at some level, over the past year. They also report that the amount of misreporting caught by internal controls in

[30]The PWC report can be seen at http://www.pwc.com/us/en/forensic-services/publications/2009-global-economic-crime-survey.jhtml.

trending down over time. Also, only 7% of misreporting and fraud cases were initially discovered by whistleblowing-related activity. In addition to the uptick in financial fraud reported in the PWC survey, the SEC Enforcement Division announced that in 2011 it filed the most enforcement actions ever in a single year. [31] The evidence suggests that, despite the massive amount of resources spent on fraud prevention in the past ten years, fraud and misreporting are no less pervasive than they were before.

Statement on Auditing Standards 99, issued by the Auditing Standards Board of the American Institute of Certified Public Accountants (AICPA) in October 2002, describes the fraud triangle. The fraud triangle is used by auditors to assess fraud risk because, generally, the three fraud triangle conditions (incentive, opportunity, and rationalization) are present when fraud occurs. First, there is an incentive or pressure that provides a reason to commit fraud. Second, there is an opportunity, and ability, for fraud to be perpetrated (e.g., absence of controls, ineffective controls, or the ability of management to override controls.) Third, the individuals committing the fraud possess an attitude that enables them to rationalize the fraud.

Hogan et al. (2008) suggest that, unlike incentive and opportunity, rationalization has received little attention from researchers. This research posits that monitoring affects an individual's attitude toward misreporting. Attitude is highly correlated with intent (Ajzen and Fishbein 2005), which triggers action and rationalization. Rationalization is described by Tsang (2002) as the cognitive process that individuals use to convince themselves that their behavior does not violate their moral standards. The way the fraud

[31]The SEC press release is found at http://www.sec.gov/news/press/2011/2011-234.htm.

triangle is conceptualized may need to be re-evaluated if it is shown that increased monitoring makes it easier to rationalize misreporting.

The standard assumption of the fraud triangle is that incentives and pressure motivate misreporting while lax controls facilitate misreporting (Hogan et al. 2008). Individuals are generally viewed as being predisposed to character traits that partially determine the extent to which they rationalize their deviant behavior (Murphy 2012). However, the theory proposed in this paper is that not only do incentives promote misreporting but strong controls may also promote deviant behavior by crowding out the intrinsic motivation to be honest, making it easier to rationalize dishonest behavior. A similar line of thought was explored by Belot and Schröder (2013). In their research experiment participants were hired for a job which had several options for deviant behavior (poor performance, tardiness, or theft). They found that increasing monitoring on one measure (performance) led to increased deviance in another measure (tardiness). They concluded that workers do "retaliate" in some way for being monitored.

The findings in this line of research have implications for the study, not only of managerial misreporting, but also whistleblowing, collusion, and worker satisfaction. As the recent wave of public accountants involved in insider trading scandals has shown, most fraud involves several collaborators—inside the firm, and sometimes outside the firm—who turn a blind eye to the unethical behavior (Burns and Kedia 2008). Often the collaborators and potential whistleblowers have different incentives, attitudes, and personality traits, but are subject to the same monitoring mechanisms. Their attitudes toward the monitoring systems may be a driver in their decisions to coalesce for or against the monitors.

HYPOTHESIS DEVELOPMENT

The fraud triangle suggests that the three conditions of incentive, opportunity, and rationalization are present when an individual commits a fraudulent act. The incentive is generally monetary in nature, while opportunity is generally conceptualized as the perception that one can perpetrate the fraud while not getting caught (Murphy and Dacin 2011). Agency theory assumes that all individuals have a natural predisposition toward fraud, and once an individual has the incentive and the opportunity to commit fraud, the rationalization is as simple as a cost-benefit calculation. However, and as psychology theories suggest, prior accounting literature has shown that individuals act more honestly than agency theory would predict (e.g., Evans et al. 2001; Hannan et al. 2006), suggesting that other influences, such as past experience and the environment, impact individuals' ability to commit and rationalize fraudulent behavior.

Researchers have identified several categories of rationalization that are often employed by perpetrators, such as moral justification, advantageous comparison, euphemistic labeling, minimization of the act, denial of the victim, and diffusion of responsibility (Murphy and Dacin 2011). However, understanding how individuals rationalize fraudulent behavior does not fully explain what characteristics of the environment, or situation, prompted the individuals to act out the deviant behavior. After all, most individuals in a position to commit fraud have a good reputation (Anand et al. 2004) which facilitates their ability to deceive others. In this study I posit that monitoring can affect an individual's attitude toward dishonesty by crowding their intrinsic motivation to be honest and increasing their ability to rationalize deviant behavior, as shown in Figure 4.1.

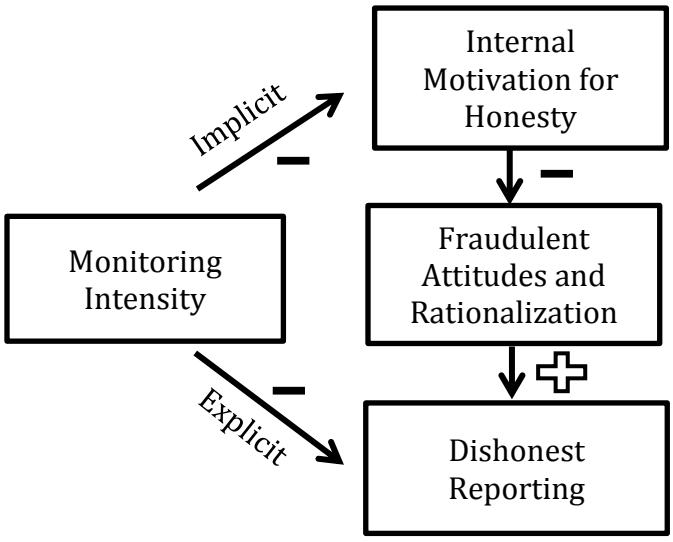


Figure 4.1 Theoretical Model

On any particular task, misreporting is directly influenced by the level of monitoring on the person reporting. For example, individuals may be inclined to cheat under 100% monitoring, but one is unlikely to cheat, misreport, or be dishonest if they know for certain they will be caught. However, not misreporting does not mean that one does not have an inclination to misreport. This inclination may be an important factor in the decision making process when the opportunity to cheat arises.

Trust and reciprocity have been widely studied in the economics literature (see Fehr and Gächter 1998). One robust conclusion from this research is that when individuals are trusted they reciprocate with trustworthy behavior (for example, Fehr et al. 1993; Berg et al. 1995; McCabe et al. 2003). Conversely, research on monitoring and surveillance has shown that individuals view monitoring, under certain conditions, as a

signal of distrust[32] (Cialdini 1996; Falk and Kosfeld 2006). This signal of distrust may lower the agent's internal motivation to treat the principal fairly and increase the agent's ability to rationalize dishonest behavior. Based on this logic I derive the following hypothesis:

> Hypothesis: When controls are removed or weakened, dishonest behavior will be higher in an environment where monitoring intensity is higher.

RESEARCH DESIGN

This experiment was carried out in a computer lab at a large public university. Using a 3X2 experimental design, where each cell included 19 participants, each of the 114 participants was subjected to one of three monitoring treatments and one of two reporting regimes. Each of the six treatment combinations included two sessions, for a total of 12 research sessions, with each session including either 9 or 10 participants. The session dates and times were pre-assigned, and participants self-registered online for the session they preferred.

Participants were recruited through the university email newsletter, flyers, and word of mouth. A diverse group 114 adult volunteers participated. Table 4.1 shows the demographics collected from the participants with a short demographics questionnaire, as shown in Appendix A, given upon arrival.

[32] The signaling of trust and distrust is important in many domains. For example, Mahar (2003) discusses how many people do not show interest in prenuptial agreements because they do not want to signal distrust in the pre-marital relationship.

Table 4.1
Demographics of Research Sample by Treatment Group and Reporting Regime

Treatment	Trust Monitored		Human Monitored		Electronically Monitored			
Reporting Regime	Self	Verified	Self	Verified	Self	Verified	**Total**	**Percent**
Gender								
Male	9	10	9	8	8	10	54	47%
Female	10	9	10	11	11	9	60	53%
							114	100%
Age								
18-24	12	11	11	15	7	11	67	59%
Over 24	7	8	8	4	12	8	47	41%
							114	100%
Student Nationality								
Domestic	8	15	13	11	9	12	68	60%
International	11	3	3	7	8	4	36	32%
Not a Student	0	1	3	1	2	3	10	9%
							114	100%
Business Student								
Business Student	9	5	6	5	8	3	36	32%
Non-Business Student	10	13	10	13	8	13	67	59%
Not a Student	0	1	3	1	3	3	11	10%
							114	100%
College Level								
Fresh/Soph	9	10	6	12	3	5	45	39%
Junior/Senior	7	6	6	2	7	5	33	29%
Graduate	3	2	4	4	6	4	23	20%
Non Student	0	1	3	1	3	5	13	11%
							114	100%
Mental Math								
Yes	9	14	15	14	10	10	72	63%
No	10	5	4	5	9	9	42	37%
							114	100%
Tired								
Yes	7	9	5	8	8	6	43	38%
No	12	10	14	11	11	13	71	62%
							114	100%

Note: Each of the 6 combinations of monitoring treatment and reporting regime had 19 participants. The Mental Math category shows the answer to the question, I consider myself good with mental math and numbers. While the Tired category shows the answer to the question, I feel tired today.

After administering the demographic questionnaire, the researcher explained the work schedule and compensation for the participants. The participants in this experiment had just spent an hour performing two distractor tasks in other experiments not related to the honesty test. These tasks served to accustom the participants to the environment, induce the monitoring treatment, and conceal the fact that their honesty was being tested. In the first distractor task, participants spent about 27 minutes performing a clerical task where they corrected data in a spreadsheet, for a flat $10 wage. In the second distractor task the participants spent about 24 minutes solving logic puzzles for a piecewise wage up to $3. The task that tested their honesty in the current study is explained in detail below.

The monitoring treatments were the same ones to which the participants had become accustomed. Individuals in the electronic monitoring treatment had one small webcam facing their workstation keyboard and papers when they arrived, as shown in Appendix B. The electronically monitored group received all the same task instructions as the other two groups, except they were told that "you are being monitored with webcams so we can observe your work and make sure you follow the instructions as given."[33] Unbeknownst to the participants, the webcams were not activated. However, the experimenter and research assistant sat prominently at a corner workstation, which the subjects believed was the "monitoring station," while all the tasks were completed.

Participants in the human monitored group were subjected to traditional human monitoring and received all the same task instructions as the other two groups, except they were told that "I will walk around the room so I can observe your work and make

[33] This wording used in the instructions is based on the wording used in a similar study by Enzle and Anderson (1993). Their study looked at the effects of controlling versus non-controlling electronic monitoring on intrinsic motivation.

sure you follow the instructions as given." The researcher and the research assistant wandered around the room and passively observed the participants during all the tasks.

Individuals in the trust treatment received all the same task instructions as the other groups. However, they were told that "you will not be watched and we believe you will follow instructions as given." All research personnel then left the room and returned when time was up for each task.

The task for this experiment was a short math puzzle. Following Ariely et al. (2009), participants were given a sheet of paper with 20 numeric matrices, as shown in Appendix H. Each matrix contained 12 three-digit numbers (e.g., 4.29, 3.23). Participants had 5 minutes to find the unique two numbers that add to 10.00 in as many matrices as possible. Participants were told, truthfully, that all the matrices had the unique combination and that they could work the sheet in any manner or order they like. Also, It was explained that they could earn anywhere from $0 to $5 on this task, depending upon their performance.

Half the participants in each monitoring treatment were told that writing or marking on the paper during the work was optional, and no indication or proof that the combination was actually found would be required. This half of the participants self-reported the number of matrices solved and their work was not verified. Thus, cheating without detection was possible. The other half of the participants in each monitoring treatment had to mark their papers to indicate the correct combination, and their work was verified. Dishonesty was operationalized as the difference in the mean scores between the self-score and non-self-score participants within a monitoring treatment group.

RESULTS

Table 4.2 shows the number of matrices reported as solved across monitoring treatment and reporting regimes. In the trust treatment, the participants who self-reported their results reported solving fewer matrixes (8.26) than the participants who knew their work would be verified (9.68). In the human monitoring treatment, the participants who self-reported their results reported solving more matrixes (11.58) than the participants who knew their work would be verified (8.47). Similarly, in the electronic monitoring treatment, participants who self-reported their results reported solving more matrixes (8.95) than the participants who knew their work would be verified (8.11).

Table 4.2
Matrices Reported Solved by Treatment and Reporting Regime

Treatment	Trust Monitored			Human Monitored			Electronically Monitored			
Reporting Regime	Self	Verified	Combined	Self	Verified	Combined	Self	Verified	Combined	ALL
Average Matrices	8.26	9.68	8.97	11.58	8.47	10.02	8.95	8.11	8.53	9.18
Std Dev	4.87	4.73	4.79	4.07	4.88	4.70	3.91	5.13	4.52	4.67
Min	0	2	0	3	1	1	4	2	2	0
Max	20	18	20	20	17	20	20	20	10	20
Obs	19	19	38	19	19	38	19	19	38	114

Note: This table shows the number of matrices reported solved for each combination of monitoring and reporting regime.

Before proceeding to the formal testing of the hypothesis it is important to note that there were three extraneous individual differences, from the demographic survey (shown in Appendix A), which were found to significantly affect the number of matrices reported as solved ($\alpha = .10$, untabulated). On average, males, those who said that they were good at mental math, and those who said that they were not tired, reported that they solved more matrices than females, those who said they were not good with mental math, and those who said they were tired. Past research suggests that these individual differences may have a direct effect on task performance, or interact with the treatments to alter performance (or reported performance). For example, some research suggests that, on average, males are slightly better at mental math (Hyde and Mertz. 2009), but some research also suggests that males are more likely to be dishonest about their performance (Dreber and Johannesson 2008). Being tired may affect performance, but past research also has shown that individual may be more dishonest about their performance when they are tired (Ariely et al. 2009). Moreover, tired individuals may feel more pressure to perform in the presence of monitoring, causing an interaction with the monitoring treatment. Lastly, past research has shown that monitoring intensity (through work-related stress) may interact with mental ability to affect performance on tasks (Schultz and Searleman 1998).

These extraneous individual differences (mental math ability, tiredness, and gender) should be included in the analysis to reduce error variance. As a result of the quasi-randomization of participants, cell sizes, after inclusion of the controls, are sufficient to calculate the main effects and all 2- and 3-way interaction terms.[34] Levene's

[34] "Higher-order interactions occur rarely" and are difficult to interpret (van Belle 2002, 135).

test indicates that the assumption of equality of error variance is not violated ($p = .17$), reducing concerns about differences in cell sizes (Neter et al. 1990).

Table 4.3 shows that, consistent with the hypothesis, after accounting for all the control variables and interactions, the monitoring treatment and reporting regime interact to affect the number of matrices individuals reportedly solved ($p = .027$). Further, this interaction is not affected by the other control variables (none of the three-way interactions including it are significant), so that I can examine this relationship without qualification. The other significant effects and interactions in the ANOVA do not relate to my hypothesis, and are included only to control for extraneous variance in the factorial design.

Table 4.3
ANOVA Results

Number of Observations	114	**R- squared**	0.53	
		Adj R-squared	0.27	

Source	Partial SS	df	MS	F	Prob > F
Model	1318.71	41	32.16	2.01	**0.005**
Monitoring	91.38	2	45.65	2.86	*0.064*
Reporting	0.06	1	0.06	0.00	0.953
Tired	95.25	2	95.25	5.97	**0.017**
Gender	121.01	2	121.01	7.58	**0.007**
Mental Math	14.80	2	14.80	0.93	0.339
Monitoring*Reporting	121.41	2	60.70	3.80	**0.027**
Monitoring*Tired	93.10	2	46.55	2.92	*0.061*
Monitoring*Gender	57.10	2	28.55	1.79	0.175
Monitoring*Mental Math	187.67	2	93.84	5.88	**0.004**
Reporting*Tired	3.72	1	3.72	0.23	0.631
Reporting*Gender	0.12	1	0.12	0.01	0.933
Reporting*Mental Math	8.03	1	8.03	0.50	0.481
Tired*Gender	6.88	1	6.88	0.43	0.514
Tired*Mental Math	12.86	1	12.86	0.81	0.373
Gender *Mental Math	20.99	1	20.99	1.31	0.255
Monitoring*Reporting*Tired	5.76	2	2.88	0.18	0.835
Monitoring*Reporting*Gender	18.19	2	9.10	0.57	0.568
Monitoring*Reporting*Mental Math	36.07	2	18.04	1.13	0.329
Monitoring*Tired*Gender	3.31	2	1.66	0.10	0.902
Monitoring*Tired*Mental Math	2.64	1	2.64	0.17	0.686
Monitoring*Gender*Mental Math	85.61	2	42.80	2.68	*0.075*
Reporting*Tired*Gender	0.12	1	0.12	0.01	0.930
Reporting*Tired*Mental Math	74.67	1	74.67	4.68	**0.034**
Reporting*Gender*Mental Math	0.06	1	0.06	0.00	0.953
Tired *Gender* Mental Math	4.42	1	4.42	0.28	0.601
Residual	1149.78	72	15.97		
Total	2468.49	113			

Note: This table shows the five-way ANOVA results for the effects of **Monitoring** treatment (Trust, Human Monitoring, or Electronic Monitoring), **Reporting** regime (self-report or verified), and the dichotomous control variables self-assessed **Tiredness**, **Gender**, and self-assessed **Mental Math** ability on the number of **Matrices** the participants reported as solved. The **Monitoring*Reporting** interaction is the key effect of interest, and is unaffected by the control variables, as shown by the nonsignificant 3-way interactions. P-values < .05 are **bolded**; those < .10 are *italicized*.

Figure 4.2 shows the adjusted means graph for each of the treatment groups. The slope of each line indicates the effect of verification on reported scores, which is my proxy for cheating. The slopes of the lines indicate that cheating may have been present in the human-monitored treatment, and to a lesser extent in the electronically monitored treatment. No cheating is apparent in the trust monitored treatment.

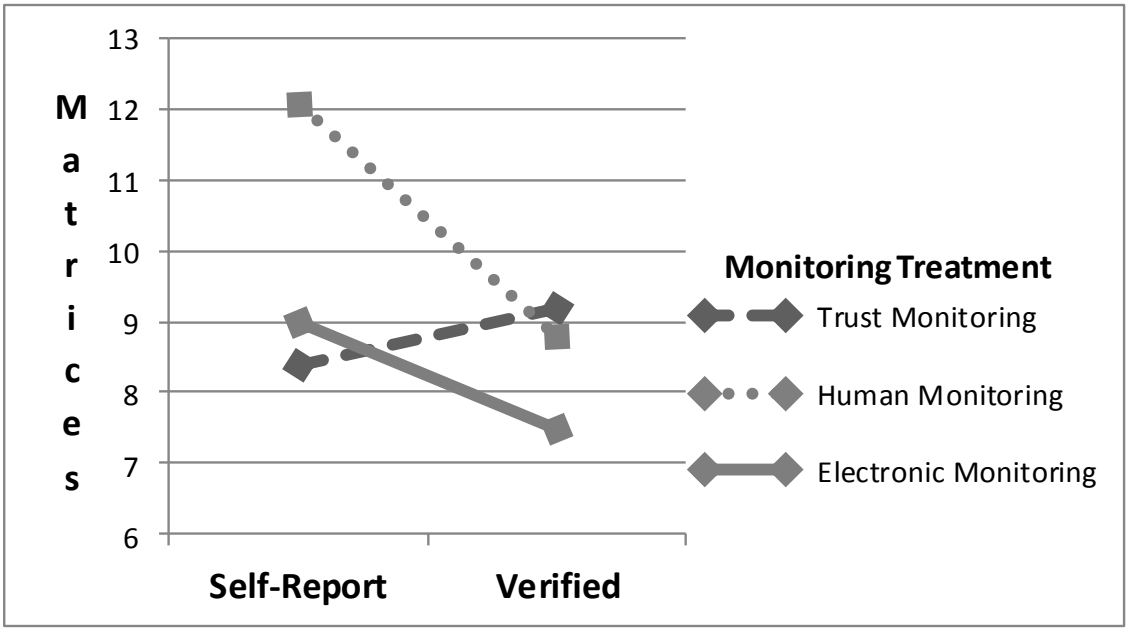


Figure 4.2: Matrices Reported as Solved: Results by Reporting Regime and Monitoring Treatment

Note: Participants either were allowed to self-report without verification, or their reports were documented and verified. Means are adjusted for Tired, Gender, and Mental Math, as reported in Table 4.4.

As shown in Table 4.4, a comparisons of the adjusted means (shown in Figure 4.2) reveals that the reporting regime treatments are not significantly different in the trust monitoring treatment (p= .593). However, the reporting regime treatment means are significantly different in the human monitored treatment (p= .035), and, while the graph does seem to indicate that cheating may have been present in the electronic monitoring treatment, the means are not significantly different at a high level of confidence (p= .275). [35] Overall, these results support my hypothesis that, given an opportunity to cheat, dishonesty will be higher in an environment where monitoring intensity is higher.

Table 4.4
Pairwise Comparison of Adjusted Means

	Reporting Regime					95% Confidence Level for Difference	
Monitoring Treatment	**Self-Report**	**Verified**	**Difference**	**Std. Error**	**Sig.**	**Lower**	**Upper**
Trust Monitoring	8.38	9.14	-0.76	1.42	0.593	-3.59	2.06
Human Monitoring	12.04	8.71	3.33	1.55	**0.035**	0.239	6.42
Electronic Monitoring	9.04	7.40	1.64	1.50	0.275	-1.34	4.63

Note: This table shows the pairwise comparison of the adjusted means for each reporting regime in each monitoring treatment. All of the comparisons were related to a planned, specific hypothesis rather than a result of post hoc comparisons. Consequently, the alpha level was not adjusted for the multiple comparisons. P-values < .05 are ***bolded.***

[35] All of the measured outcomes were related to a planned, specific hypothesis rather than a result of post hoc comparisons. For this reasons, I did not adjust the significance level of the *p* values for the planned comparisons. This approach is consistent with guidelines for planned multiple comparisons (Fisher 1947; Rothman 1990).

CONCLUSION

In this study, I theorized that monitoring could crowd-out an individual's intrinsic motivation to be honest. I further theorized that this loss of intrinsic motivation would change the individual's attitude toward dishonesty and increase their ability to rationalize deviant behavior—all leading to a higher propensity toward dishonest behavior. This led to my hypothesis that, given an opportunity to cheat, dishonesty will be higher in an environment where monitoring intensity is higher.

To test my hypothesis, I assigned each participant in the experiment to one of three monitoring environments: trust monitoring, human monitoring, or electronic monitoring. With this treatment induced, I gave the participants a simple mental math task with a monetary reward based on performance. Half the participants in each treatment self-reported their results; while the other half had their results verified (groups were segregated and unaware of each other). The spread between the average reported performance of verified and non-verified groups was used a proxy for the incidence of cheating in each monitoring-treatment group (Ariely et al. 2009).

Dishonesty was not detected in the trust treatment, but cheating was detected in the human monitored treatment and—to a lesser extent—in the electronically monitored treatment. Therefore it appears that monitoring does affect the agent's preference for honesty. Thus, I find evidence to support of my hypothesis, although questions still remain as to why cheating was detected in the human monitored treatment at a statistically significant level, but cheating was not detected at a significant level in the electronically monitored treatment.

I offer three explanations for why cheating was higher in the human monitored treatment than the electronically monitored treatment. First, it is possible that there was a strong propensity to be dishonest in the electronically monitored treatment, but the electronic monitoring convinced the participants that the risk of exposure was still present in this situation. Perhaps they feared they were being recorded, or their movements on the mental math task were being scrutinized. If so, then it is probable that, even though they had a high propensity to be dishonest, they thought it better to be honest and not risk detection. Second, it is possible that the individuals saw the electronic monitoring as a cue that the task was very important to the monitor, or that the monitor was very concerned with their work. If individuals viewed the task as being important to someone then that may have decreased their propensity to be dishonest, even if they disliked the monitoring. Third, it is possible that the participants did not dislike the electronic monitoring as much as they disliked the human monitoring, leading to lower propensity to be dishonest. This explanation would not be consistent with past research and anecdotal evidence which shows electronic monitoring is more stressful than traditional human monitoring (Stanton 2000). Future research should be done in this area to determine the how individuals view different monitoring regimes, and how their views shape their attitudes towards different work behaviors.

In conclusion, the agency theory literature and fraud prevention literature rarely consider the negative effects of monitoring on the individual psyche. Usually, only the principal's explicit monetary costs are considered when searching for the optimal amount of monitoring. This study, and others, suggest that there are significant "hidden costs" (Falk and Kosfeld 2006), and unanticipated effects, of monitoring and control that have

yet to be fully explored in the business literature. Since these costs and effects are mostly unknown, business researchers currently lack the ability to predict the effects of controls, or regulation, on behavior. In contrast, much more is known about other environmental effects on behavior, such as the effects of incentives on work performance (Bonner et al. 2000), than is known about how individuals react to different types of internal controls. Following Christ et al. (2012), I believe that future research should further develop our understanding of the potential consequences of formal controls.

Further developing this line of research may yield important clues to long-standing questions, such as why financial fraud is still persistent despite increases in regulation and ethics training (Rezaee 2005), why individuals display trustworthy behavior in certain situations and contractual arrangements but selfish behavior in others (Rankin et al. 2008), why individuals collude against control systems (Zhang 2008), and finally, why whistleblowing may be more likely in some environments or situations than others (Seifert et al. 2013). Using empirical evidence to address the questions will aid in the design of more effective internal controls (Sprinkle 2003), the development of more efficient contracts (Brown et al. 2009), and more comprehensive theoretical models for business researchers (e.g., Tirole 2009)

CHAPTER 5

CONCLUSION

The standard principal-agent theory of the firm suggests the principal should seek to control the agent's behavior (solve the agency problem) though either monitoring the agent's performance or aligning the agent's interest with the owner's interest through the use of incentives.[36] While there has been a wealth of interesting research on the effects of incentives on agent behavior (Bonner et al. 2000), less is known about the effects monitoring and control on the agent. This has led some to call for research that further develops our understanding of the potential consequences of formal controls (e.g., Christ et al. 2012). This three paper dissertation heeds this call.

Drawing on past literature, which suggests that agents have a disdain for control (Falk and Kosfield 2006) and find monitoring stressful (Stanton 2000), I formed hypotheses about how and why agents may react to various levels of monitoring. In short, the literature seems to suggest that, when all else is equal, monitoring should increase external motivation at the expense of internal motivation and reciprocity. This basic assumption motivated all three studies in this dissertation. The first study in this dissertation investigates the effects of monitoring intensity on discretionary effort, volunteer rates for an optional task, and effort spent on an optional task. The second study investigates the effects monitoring intensity on problem solving ability. The third study investigates the effects of monitoring intensity on behavioral honesty.

[36] In practice, most business arrangements seem to use a mix of incentive pay and monitoring. For example, Bulow and Summers's (1986) model of dual labor markets is based on the assumption that if monitoring is difficult work conditions will be good and pay will be high, while if monitoring is easy, working conditions will be poor and pay will be low.

Overall the results from the three studies seem to support the theory that increased monitoring intensity lowers internal motivation, although the results vary on different measures across the different types of monitoring environments (human monitoring or electronic monitoring). For example, in the first study I expected to find that monitoring intensity would be associated with lower discretionary effort, lower volunteer rates for optional tasks, and lower effort on optional tasks. While this is assumption was mostly correct, there does seem to be a more complicated relationship between volunteer rates for an optional task and monitoring intensity. It seems that in some cases high monitoring intensity causes volunteer rates for optional tasks to increase, perhaps because of a perceived loss of autonomy. In study two I expected to find that monitoring intensity would be associated with lower problem solving ability across all measures. While I found this to generally be the case, I also found that an individual's confidence in their base ability interacted with the monitoring treatment to determine performance on complex problem solving.[37] In study three I expected to find that, when given the opportunity to cheat, monitoring intensity would be associated with increased dishonesty. Again, generally I found this to be case. However, I did find that monitoring at the most intense levels seems to curtail dishonest behavior.

While the topics in these studies are becoming increasingly important to researchers and practitioners, most of the ideas, methods, and psychological theories drawn on this dissertation are novel to the managerial accounting literature. The results of these experiments are not meant to be directly generalizable outside of a laboratory setting. Rather, these experiments, and their results, should be considered in the larger

[37] These findings are consistent with the theory of social facilitation, which details how individual subconsciously change their behavior in the presence of others (e.g., Zajonc 1965; Aiello and Douthitt 2001).

frameworks of organizational theory, human psychology, and current practices. Taken together, the findings from this dissertation add to our understanding of the effects organizational controls on agent behavior and abilities. These findings also open up new research questions and avenues for research.

REFERENCES

Aiello, J.R., and E.A. Douthitt. 2001. Social facilitation from Triplett to electronic performance monitoring. *Group Dynamics: Theory, Research, and Practice 5* (3): 163–180.

———, and K. Kolb. 1995. Electronic performance monitoring and social context: Impact on productivity and stress. *Journal of Applied Psychology* 80 (3): 339–353.

———, and Y. Shao. 1993. Electronic monitoring performance and stress: *The role of feedback and goal setting. Human Computer Interaction: Applications and case studies*. Amsterdam: Elsevier Science.

Ajzen, I., and M. Fishbein. 2005. The influence of attitudes on behavior. *The Handbook of Attitudes*: 173–222. Lawrence Erlbaum Associates Publishers.

Alchian, A. A., and H. Demsetz. 1972. Production, information costs, and economic organization. *The American Economic Review* 62 (5): 777–795.

Alder, G. S., and M. L. Ambrose. 2005. Towards understanding fairness judgments associated with computer performance monitoring: An integration of the feedback, justice, and monitoring research. *Human Resource Management Review* 15 (1): 43–67.

Amabile, T. M. 1996. *Creativity in Context: Update to the Social Psychology of Creativity*. Westview Press.

Anand, V., B. Ashforth, and M. Joshi. 2004. Business as usual: The acceptance and perpetuation of corruption in organizations. *Academy of Management Executive* 18 (2): 39–53.

Arce, D. G. 2007. Is agency theory self-motivating? *Economic Inquiry* 45 (4): 708–720.

Ariely, D. 2009. *Predictably irrational, revised and expanded edition: The hidden forces that shape our decisions.* Harper House.

———, D., N. Mead, L. Baumeister, F. Gino, and M. Schweitzer. 2009. Too tired to tell the truth: Self-control resource depletion and dishonesty. *Journal of Experimental Social Psychology* 45 (3): 594–597.

———, D., N. Mazar, and O. Amir. 2008. The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research* 45 (6): 633–644.

Arrow, K. J. 1985. Informational structure of the firm. *The American Economic Review* 75 (2): 303–307.

Asare, S. K., and A. M. Wright. 2004. The effectiveness of alternative risk assessment and program planning tools in a fraud setting. *Contemporary Accounting Research* 21 (2): 325–352.

Baiman, S., and B. Lewis. 1989. An experiment testing the behavioral equivalence of strategically equivalent employment contracts. *Journal of Accounting Research* 27 (1): 1–20.

Baril, C. P., B. M. Cunningham, D. R. Fordham, R. L. Gardner, and S. K. Wolcott. 1998. Critical thinking in the public accounting profession: Aptitudes and attitudes. *Journal of Accounting Education* 16 (3): 381–406.

Beccerra, M., and A. Gupta. 1999. Trust within the organization: Integrating the trust literature with Agency theory and transaction costs economics. *Public Administration Quarterly* 23 (2): 177–203.

Belot, M., and M. Schröder. 2013. Does monitoring work? A field experiment with multiple forms of counterproductive behavior. Working Paper Series University of Magdeburg.

Berg, J., J. Dickhaut, and K. McCabe. 1995. Trust, reciprocity, and social history. *Games and Economic Behavior* 10 (1): 122–142.

Bernheim, B. D., and M. D. Whinston. 1998. Incomplete contracts and strategic ambiguity. *American Economic Review* 88 (4): 902–932.

Bierstaker, J. L., J. C. Bedard, and S. F. Biggs. 1999. The role of problem representation shifts in auditor decision processes in analytical procedures. *Auditing: A Journal of Practice & Theory* 18 (1): 18–36.

————, J. L., and S. Wright. 2001. A research note concerning practical problem-solving ability as a predictor of performance in auditing tasks. *Behavioral Research in Accounting* 13 (1): 49–62.

Bilalić, M., P. McLeod, and F. Gobet. 2008. Inflexibility of experts—Reality or myth? Quantifying the Einstellung effect in chess masters. *Cognitive Psychology* 56 (2): 73–102.

Blascovich, J., W.B. Mendes, S.B. Hunter, and K. Salomon. 1999. Social facilitation as challenge and threat. *Journal of Personality and Social Psychology 77 (1):* 68–77.

Bloom, R., P. T. Elgers, and D. Murray. 1984. Functional fixation in product pricing: A comparison of individuals and groups. *Accounting, Organizations and Society* 9 (1): 1–11.

Bond, C.F., and L.J. Titus. 1983. Social facilitation: A meta-analysis of 241 studies. *Psychological Bulletin 94* (1): 265–292.

Bonner, S. E., and B. L. Lewis. 1990. Determinants of auditor expertise. *Journal of Accounting Research 28 (3)*: 1–20.

————, S., R. Hastie, G. Sprinkle, and S. Young. *2000*. A Review of the effects of financial incentives on performance in laboratory tasks: Implications for management accounting. *Journal of Management Accounting Research* 12 (1): 19–64.

Borman, W. C., and S. J. Motowidlo. 1993. Expanding the criterion domain to include elements of contextual performance. *Personnel Selection in Organizations* (79) 4: 71–98.

Brown, J., J. Evans III, and D. Moser. 2009. Agency theory and participative budgeting experiments. *Journal of Management Accounting Research* 21 (1): 317–345.

Bulow, J. I., and L.H. Summers. 1986. A Theory of Dual Labor Markets with Application to Industrial Policy, Discrimination, and Keynesian Unemployment. *Journal of Labor Economics* 4 (3): 376–414.

Burns, N., and S. Kedia. 2008. Executive option exercises and financial misreporting. *Journal of Banking & Finance* 32 (5): 845–857.

Callahan, C., and J. R. Larson. 1990. Performance monitoring: How it affects work productivity. *Journal of Applied Psychology* 75 (5): 530–538.

Carnevale, P. J., and T. M. Probst. 1998. Social values and social conflict in creative problem solving and categorization. *Journal of Personality and Social Psychology* 74 (5): 1300–1309.

Christ, M. H., S. A. Emett, S. L. Summers, and D. A. Wood. 2012. The effects of preventive and detective controls on employee performance and motivation. *Contemporary Accounting Research* 29 (2): 432–452.

Cialdini, R. B. 1996. The triple tumor structure of organizational behavior. In D. M. Messick & A. E.Tenbrunsel (Eds.), *Codes of conduct* 44–58. New York: Sage.

Cowen, E. L. 1952. The influence of varying degrees of psychological stress on problem-solving rigidity. *The Journal of Abnormal and Social Psychology* 47 (2S): 512–519.

Cressey, D. R. 1973. *Other People's Money*. Montclair: Patterson Smith.

Crutchley, C. E., M. R. H. Jensen, and B. B. Marshall. 2007. Climate for scandal: Corporate environments that contribute to accounting fraud. *Financial Review* 42 (1): 53–73.

Cuevas-Rodríguez, G., L. R. Gomez-Mejia, and R. M. Wiseman. 2012. Has Agency Theory Run its Course?: Making the Theory more Flexible to Inform the Management of Reward Systems. *Corporate Governance: An International Review* 20 (6): 526–546.

Davidson, R. and R. Henderson. 2000. Electronic performance monitoring: A Laboratory investigation of the influence of monitoring and difficulty on task performance, mood state, and self-reported stress levels. *Journal of Applied Psychology 30* (5): 906–920.

Dearman, D. T., and M. D. Shields. 2005. Avoiding accounting fixation: Determinants of cognitive adaptation to differences in accounting method. *Contemporary Accounting Research* 22 (2): 351–84.

Deci, E. L., R. Koestner, and R. M. Ryan. 1999. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin* 125 (6): 627–688.

————, E. L., R. J. Vallerand, L. G. Pelletier, and R. M. Ryan. 1991. Motivation and education: The self-determination perspective. *Educational Psychologist* 26 (3-4): 325–346.

Dickinson, D., and M. C. Villeval. 2008. Does monitoring decrease work effort? The complementarity between agency and crowding-out theories. *Games and Economic Behavior* 63 (1): 56–76.

Dreber, A., and M. Johannesson. 2008. Gender differences in deception. *Economics Letters* 99 (1): 197–199.

Eisenhardt, K. M. 1989. Agency theory: An assessment and review. *Academy of Management Review* 14 (1): 57–74.

Elsbach, K. D., and A. B. Hargadon. 2006. Enhancing creativity through "mindless" work: A framework of workday design. *Organization Science* 17 (4): 470–83.

Enzle, M. E., and S. C. Anderson. 1993. Surveillant intentions and intrinsic motivation. *Journal of Personality and Social Psychology* 64 (2): 257–266.

Erikson, J. W. 2012. Problem-solving: Individual factors predictive of resistance to functional fixedness and effects of Einstellung. Working Paper Indiana State University.

Evans III, J. H., R. L. Hannan, R. Krishnan, and D. V. Moser. 2001. Honesty in managerial reporting. *The Accounting Review* 76 (4): 537–559.

Falk, A., and M. Kosfeld. 2006. The hidden costs of control. *The American Economic Review* 96 (5): 1611–1630.

Fantino, E., B. A. Jaworski, D. A. Case, and S. Stolarz-Fantino. 2003. Rules and problem solving: Another look. *The American Journal of Psychology 116 (4): 613–632.*

Fehr, E., and S. Gächter. 1998. Reciprocity and economics: The economic implications of homo Reciprocans. *European Economic Review* 42, (3-5): 845–859.

_____, and A. Falk. 2002. Psychological foundations of incentives. *European Economic Review* 46 (4): 687–724.

_____, G. Kirchsteiger, and A. Riedl. 1993. Does fairness prevent market clearing? An experimental investigation. *The Quarterly Journal of Economics* 108 (2): 437–59.

Feist, G. J. 1998. A meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Review* 2 (4): 290–309.

Ferraro, F., J. Pfeffer, and R. I. Sutton. 2005. Economics language and assumptions: How theories can become self-fulfilling. *Academy of Management Review* 30 (1): 8–24.

Fessler, N. J. 2003. Experimental evidence on the links among monetary incentives, task attractiveness, and task performance. *Journal of Management Accounting Research* 15 (1): 161–176.

Fisher, R.A. 1947. *The Design of Experiments*. Fourth Edition. Edinburgh: Oliver and Boyd.

Frederickson, J.R., and C.B. Cloyd. 1998. The effects of performance cues, subordinate susceptibility to social influences, and the nature of the subordinate's private information on budgetary slack. *Advances in Accounting* 16 (1): 89–115.

Frey, B. S. 1993. Does monitoring increase work effort? The rivalry with trust and loyalty. *Economic Inquiry* 31 (4): 663–670.

———. 1994. How intrinsic motivation is crowded out and in. *Rationality and Society* 6 (3): 334–352.

_____, and F. Oberholzer-Gee. 1997. The cost of price incentives: An empirical analysis of motivation crowding-out. *The American Economic Review* 87 (4): 746–755.

Fulmer, I. S., B. Barry, and D. A. Long. 2009. Lying and smiling: Informational and emotional deception in negotiation. *Journal of Business Ethics* 88 (4): 691–709.

Gächter, S. and A. Falk. 2002. Reputation and reciprocity: Consequences for the labour relation. *The Scandinavian Journal of Economics* 104 (1): 1–26.

Gagné, M., and E. L. Deci. 2005. Self determination theory and work motivation. *Journal of Organizational Behavior* 26 (4): 331–362.

Gibbins, M., and K. Jamal. 1993. Problem-centered research and knowledge-based theory in the professional accounting setting. *Accounting, Organizations and Society* 18 (5): 451–66.

Gibson, R., C. Tanner, and A. Wagner. 2012. Preferences for truthfulness: Heterogeneity among and within individuals. *American Economic Review* 103 (1): 532–548.

Gneezy, U., S. Meier, and P. Rey-Biel. 2011. When and why incentives don't work to modify behavior. *The Journal of Economic Perspectives* 25 (4): 191–209.

———, U. 2005. Deception: The role of consequences. *The American Economic Review* 95 (1): 384–394.

Gould, S. J. 1996. *The mismeasure of man.* WW Norton & Company.

Hannan, R. L., F. W. Rankin, and K. L. Towry. 2006. The effect of information systems on honesty in managerial reporting: A behavioral perspective. *Contemporary Accounting Research* 23 (4): 885–918.

Hansen, S. C. (1997). Designing internal controls: The interaction between efficiency wages and monitoring. *Contemporary Accounting Research* 14(1): 129–163.

Hatcher, L., T. L. Ross, and D. Collins. 1989. Prosocial behavior, job complexity, and suggestion contribution under gainsharing plans. *The Journal of Applied Behavioral Science* 25 (3): 231–248.

Hecht, G., I. Tafkov, and K. L. Towry. 2012. Performance spillover in a multitask environment. *Contemporary Accounting Research*. 29 (2): 563–589.

Heiman, V. B. 1990. Auditors' assessments of the likelihood of error explanations in analytical review. *The Accounting Review 65 (4)*: 875–890.

Hennessey, B. A., and T. M. Amabile. 1998. Reality, intrinsic motivation, and creativity. *The American Psychologist* 53 (6): 674–675.

Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, and R. McElreath. 2001. In search of homo economicus: Behavioral experiments in 15 small-scale societies. *American Economic Review* 91 (2): 73–78.

Hogan, C. E., Z. Rezaee, R. A. Riley, Jr., and U. K. Velury. 2008. Financial statement fraud: Insights from the academic literature. *Auditing: A Journal of Practice & Theory 27 (2*): 231–252.

Hölmstrom, B. 1979. Moral hazard and observability. *The Bell Journal of Economics* 10 (1): 74–91.

————, B. and P. Milgrom. 1991. Multitask principal-agent analyses: incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization* (7): 24–52.

Hyde, J. S., and J.E. Mertz (2009). Gender, culture, and mathematics performance. *Proceedings of the National Academy of Sciences* 106 (22): 8801–8807.

Isen, A. M., K. A. Daubman, and G. P. Nowicki. 1987. Positive affect facilitates creative problem solving. *Journal of Personality and Social Psychology* 52 (6): 1122–1131.

Jensen, M. C., and W. H. Meckling. 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* 3 (4): 305–360.

Klein, C. T. F., and D. M. Webster. 2000. Individual differences in argument scrutiny as motivated by need for cognitive closure. *Basic and Applied Social Psychology* 22 (2): 119–129.

Knoblich, G., S. Ohlsson, and G. E. Raney. 2001. An eye movement study of insight problem solving. *Memory & Cognition* 29 (7): 1000–1009.

Koontz, H., and C. O'Donnell. 1955. P*rinciples of Management: An Analysis of Managerial Functions*. McGraw-Hill, New York, NY.

Kruglanski, A. W., and T. Freund. 1983. The freezing and unfreezing of lay-inferences: Effects on impressional primacy, ethnic stereotyping, and numerical anchoring. *Journal of Experimental Social Psychology* 19 (5): 448–468.

Kuang, X., and D. V. Moser. 2009. Reciprocity and the effectiveness of optimal agency contracts. *The Accounting Review* 84 (5): 1671–1694.

Kulik, B.W. 2005. Agency theory, reasoning and culture at Enron: In search of a solution. *Journal of Business Ethics* 59 (4): 347–360.

Lambert, R. A. 2001. Contracting theory and accounting. *Journal of Accounting and Economics* 32 (1): 3–87.

Libby, R., and J. Luft. 1993. Determinants of judgment performance in accounting settings: Ability, knowledge, motivation, and environment. *Accounting, Organizations and Society* 18 (5): 425–450.

Luchins, A. S. 1942. Mechanization in problem solving: The effect of Einstellung. *Psychological Monographs* 54 (6): 1–95.

Lund, J. 1992. Electronic performance monitoring: A review of research issues. *Applied Ergonomics* 23 (1): 54–58.

Lundquist, T., T. Ellingsen, E. Gribbe, and M. Johannesson. 2007. The cost of lying. *Working Paper Series in Economics and Finance*, Stockholm School of Economics, No. 666.

Mahar, H. 2003. Why are there so few prenuptial agreements? *Harvard Law School John M.Olin Center for Law, Economics and Business Discussion Paper Series*: 436.

Marchant, G., J. Robinson, U. Anderson, and M. Schadewald. 1991. Analogical transfer and expertise in legal reasoning. *Organizational Behavior and Human Decision Processes* 48 (2): 272–290.

Matuszewski, L. J. 2010. Honesty in managerial reporting: Is it affected by perceptions of horizontal equity? *Journal of Management Accounting Research* 22 (1): 233–250.

McCabe, K. A., M. L. Rigdon, and V. L. Smith. 2003. Positive reciprocity and intentions in trust games. *Journal of Economic Behavior & Organization* 52 (2): 267–275.

Mock, T. J., and A. Wright. 1993. An exploratory study of auditors' evidential planning judgments. *Auditing* 12 (2): 39–68.

Motowildo, S. J., and J. R. Van Scotter. 1994. Evidence that task performance should be distinguished from contextual performance. *Journal of Applied Psychology* 79 (4): 475–480.

Murphy, P. R. 2012. Attitude, Machiavellianism and the rationalization of misreporting. *Accounting, Organizations and Society* 37 (4): 242–259.

————, and M. T. Dacin. 2011. Psychological pathways to fraud: Understanding and preventing fraud in organizations. *Journal of Business Ethics* 101 (4): 601–618.

Neter, J., W. Wasserman, and M. H. Kutner. 1990. *Applied Linear Statistical Models*, 3rd ed., Irwin.

Pendergast, C. 2000. What trade-off of risk and incentives? *The American Economic Review* 90 (2): 421–455.

Public Company Accounting Oversight Board (PCAOB) (2005). Consideration of fraud in a financial statement audit. AU Section 316.

Ramamoorti, S. 2008. The psychology and sociology of fraud: Integrating the behavioral sciences component into fraud and forensic accounting curricula. *Issues in Accounting Education* 23 (4): 521–533.

Rankin, F. W., S. T. Schwartz, and R. A. Young. 2008. The effect of honesty and superior authority on budget proposals. *The Accounting Review* 83 (4): 1083–1099.

Rezaee, Z. 2005. Causes, consequences, and deterrence of financial statement fraud. *Critical Perspectives on Accounting* 16 (3): 277–298.

Rothman, K. J. 1990. No adjustments are needed for multiple comparisons. *Epidemiology* 1 (1): 43–46.

Rousseau, D., S.B. Sitkin, R.S. Burt, and C. Camerer. 1998. Not So Different After All: A Cross-Discipline View of Trust. *Academy of Management Review* 23 (3): 393–404.

Ryan, R. M., and E. L. Deci. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist* 55, (1): 68–78.

Salterio, S. E., and A. Webb. 2006. Honesty in accounting and control: A discussion of "The effect of information systems on honesty in managerial reporting: A behavioral perspective." *Contemporary Accounting Research* 23 (4): 919–932.

Schultz, P., and A. Searleman. 1998. Personal need for structure, the Einstellung task, and the effects of stress. *Personality and Individual Differences* 24 (3): 305–310.

———, W. F. Stone, and R. Christie. 1997. Authoritarianism and mental rigidity: The Einstellung problem revisited. *Personality and Social Psychology Bulletin* 23, (1): 3–9.

Seifert D.L., W. Stammerjohan, and R. B. Martin. 2013. Trust, Organizational Justice, and Whistleblowing: A Research Note. *Behavioral Research in Accounting* (forthcoming).

Shalley, C. E., J. Zhou, and G. R. Oldham. 2004. The effects of personal and contextual characteristics on creativity: Where should we go from here? *Journal of Management* 30 (6): 933–958.

Smith, S. M., and S. E. Blankenship. 1991. Incubation and the persistence of fixation in problem solving. *The American Journal of Psychology 104 (1)*: 61–87.

Smith, M., P. Carayon, K. Sanders, S. Lim, and D. Legrande. Employee stress and health complaints in jobs with and without electronic performance monitoring. *Applied Ergonomics* 23 (1): 17–27.

Spector, P.E. 1986. Perceived control by employees: A meta-analysis of studies concerning autonomy and participation at work. *Human Relations* 39 (11): 1005–1016.

Sprinkle, G. B. 2003. Perspectives in experimental research in managerial accounting. *Accounting, Organizations and Society* 28 (2–3): 287–318.

Stanton, J. M. 2000. Reactions to employee performance monitoring: Framework, review, and research directions. *Human Performance* 13 (1): 85–113.

————, J. M., and S. T. M. Sarkar-Barney. 2003. A detailed analysis of task performance with and without computer monitoring. *International Journal of Human-Computer Interaction* 16 (2): 345–366.

————, J. M., and E. M. Weiss. 2000. Electronic monitoring in their own words: An exploratory study of employees' experiences with new types of surveillance. *Computers in Human Behavior* 16 (4): 423–440.

Stedry, A. C. 1960. *Budget Control and Cost Behavior*. The Ford Foundation Doctoral Dissertation Series 1959 Award Winner. Englewood Cliffs, NJ: Prentice Hall.

Stevens, D. E., and A. Thevaranjan. 2010. A moral solution to the moral hazard problem. *Accounting, Organizations and Society* 35 (1): 125–139.

Tayler, W. B., and R. J. Bloomfield. 2011. Norms, conformity, and controls. *Journal of Accounting Research* 49 (3): 753–790.

Tierney, P., and S. M. Farmer. 2004. The pygmalion process and employee creativity. *Journal of Management* 30 (3): 413–432.

Tirole, J. 2009. Cognition and Incomplete Contracts. *American Economic Review* 99 (1): 265–294.

Tosi, H. L., J.P. Katz, and L.R. Gomez-Mejia. 1997. Disaggregating the agency contract: The effects of monitoring, incentive alignment, and term in office on agent decision making. *Academy of Management Journal*, 40 (3): 584–602.

Tsang, J. A. 2002. Moral rationalization and the integration of situational factors and psychological processes in immoral behavior. *Review of General Psychology* 6 (1): 25–50.

Utman, C. H. 1997. Performance effects of motivational state: A meta-analysis. *Personality and Social Psychology Review* 1 (2): 170–182.

van Belle, G. 2002. *Statistical Rules of Thumb.* Wiley-Interscience.

Wiersma, U. J. 2011. The effects of extrinsic rewards in intrinsic motivation: A meta-analysis. *Journal of Occupational and Organizational Psychology* 65 (2): 101–114.

Wiley, J. 1998. Expertise as mental set: The effects of domain knowledge in creative problem solving. *Memory & Cognition* 26 (4): 716–730.

Wilner, N., and J. Birnberg. 1986. Methodological problems in functional fixation research: Criticism and suggestions. *Accounting, Organizations and Society* 11 (1): 71–80.

Wright, A. M., and J. C. Bedard. 2000. Decision processes in audit evidential planning: A multistage investigation. *Auditing: A Journal of Practice & Theory* 19 (1): 123–43.

Zajac, E. and J. Westphal. 1994. The costs and benefits of managerial incentives and monitoring in large U.S. corporations: When is more not better? *Strategic Management Journal* 15 (S-1): 121–142.

Zajonc, R.B. 1965. Social facilitation. *Science* 149 (1): 269–274.

———, 1980. Comprescence. In P.B. Paulus (Ed.) *Psychology of Group Influence.* Hillsdale, New Jersey: Erlbaum.

Zhang, Y. 2008. The effects of perceived fairness and communication on honesty and collusion in a multi-agent setting. *The Accounting Review* 83 (4): 1125–1.

Zhou, J., and C. E. Shalley. 2003. Research on employee creativity: A critical review and directions for future research. *Research in Personnel and Human Resources Management* 22: 165–217.

Appendix A
Short Demographics Questionnaire

Circle all that apply to you

1) I am: **Male Female**

2) My age is : **Under 18 18 -24 Over 24**

3) I consider myself mostly an: **International Student American Student Not a student**

4) I consider myself mostly: **A business student Not a business student Not a student**

5) I am a: **Freshman/Sophomore Junior/Senior Graduate Student Other**

6) I like to play sports or enjoy watching sports: **Yes No**

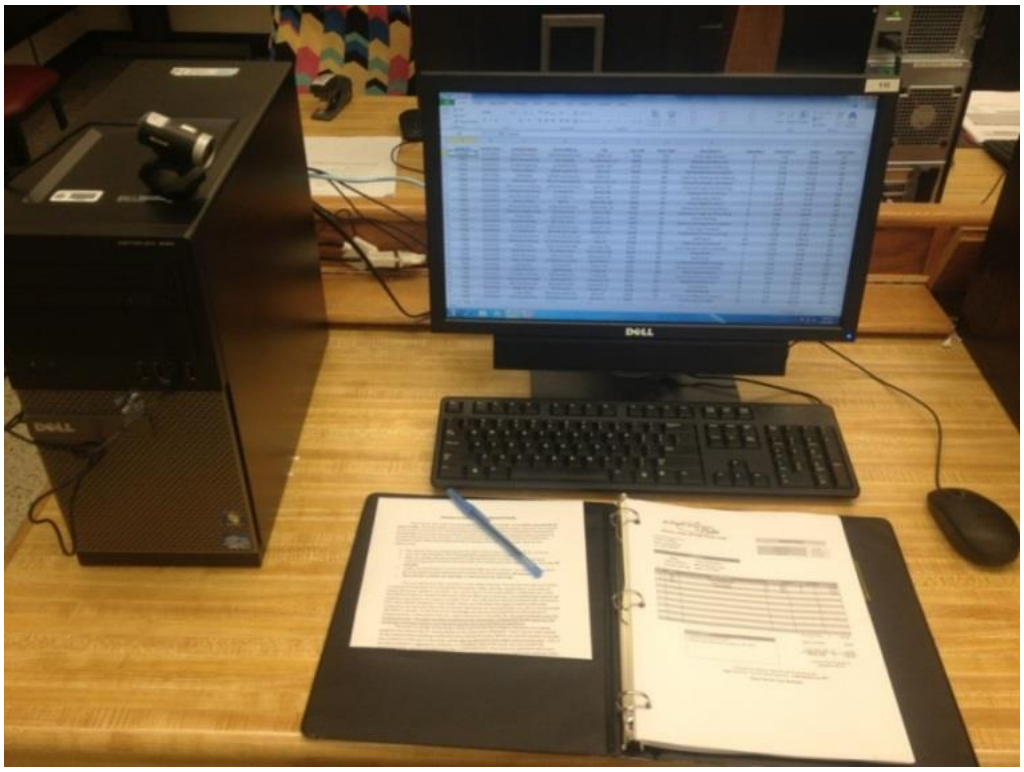
7) I consider myself good with numbers and mental math: **Yes No**

8) I am in a good mood (happy) today: **Yes No**

9) I am tired today: **Yes No**

Your answers on this form and your performance on the assigned tasks will remain anonymous and will not be matched to your name, image, person, or consent form

Appendix B
Electronic Monitoring Workstation with Camera on Computer Tower



Appendix C
Example of Task Invoice

education USA

America's Math and Logic Puzzle Leader

Invoice Due

1721 N Germantown Pkwy
Cordova, TN 38016
Phone: 1-800-296-0673
Fax: 1-901-296-0678

Invoice #	77613
Invoice Date	12/14/2012
Credit Terms	2/15, N/30

Customer

Name: Tiller Homeschool
Street Address: 116 Craig RD
City and Zip Code: Ft. Johnson, CO 80055

Item	SKU	DESCRIPTION	Price Each	Quantity	AMOUNT
1	61D	Weights and Measures	$ 5.49	6	32.94
2	16C	Power Reading	$ 4.99	9	44.91
3	33E	Audio Books	$ 12.99	2	25.98
					-
					-
					-
					-
					-
					-
					-
					-
					-
					-
					-
					-

Pre-Tax Total	$	103.83
Sales Tax Rate		6.25%
Total Sales Tax	$	6.49
TOTAL DUE	$	110.32

OTHER COMMENTS

1. Total payment due in 30 days
2. Please include the invoice number on your check

Make all checks payable to
Education USA Inc.

If you have any questions about this invoice, please contact
Reggie Thomason - Customer Service Specialist - **1-800-296-0673 Ext. 8871**

Thank You For Your Business!

Appendix D
Example of Electronically Entered Data

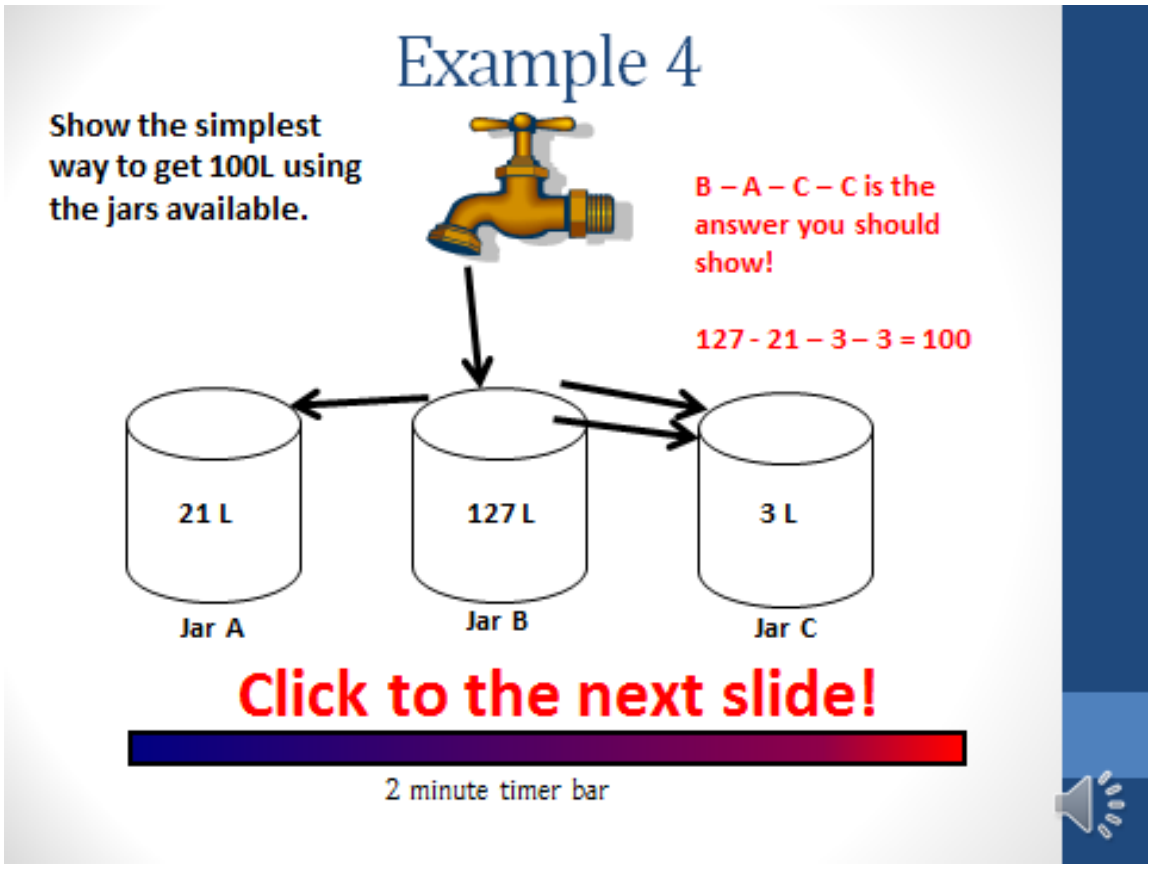
Invoice #	Invoice Date	Customer Name	Street Address	City	Zip Code	Item 1 SKU	Description 1
Sample	12/13/2012	Davis Elmentary	3402 Honeycutt Lane	Salem, AL	35444	72A	Fuzzy Logic Concept
77613	12/14/2012	Tiller Homeschool	116 Craig RD	Ft. Johntson, CO	80055	61D	Weights and Measures
77614	12/14/2012	Gradar Academy	21 Shallow Hill	Casper, WY	82717	11A	Making Cents of Money
77616	12/15/2012	Hoover Baptist	3636 N. Applin St.	Gary, IN	46335	31S	Elementary Numerical Basics
77617	12/15/2012	Dallas East	303030 Caraway Ave.	Dallas, TX	77231	80D	Graphing, Charting, and Mapping
77618	12/15/2012	Sherry Henson	9911 W. Harrwood St.	Junction City, SC	29563	07A	Geometry, Shapes, and Spaces
77619	12/15/2012	Mathnasium of Boston	17454 Hickory Hill Ste. C	Boston, MS	02103	63A	Advanced Logic and Reasoning
77621	12/16/2012	Chambers and Sloan	71 Jersey Cir.	Searcy, AR	72475	11A	Making Cents of Money
77622	12/16/2012	Chambers and Sloan	71 Jersey Cir.	Searcy, AR	74275	31S	Elementary Basics
77623	12/17/2012	Danny Fielder	403 Dr.	Batesville, MS	38606	39D	Plane Trigonometry II
77624	12/17/2012	Hillcrest Elementary	1324 Cresmont Dr.	Greenville, LA	70856	3S1	Elementary Numerical Basics
77625	12/17/2012	Sherman Heights Prep	321 Dover Ave.	Sistern, MA	05562	63A	Advances In Logic and Reasoning
77627	12/17/2012	Cindy Shultz	112 East Bark St.	Oakley, CA	90028	16C	Power Reading
77629	12/18/2012	Amanda Milligan	245 Lowcust Dr.	Caraway, AR	72419	22B	Fractions and Algebraic Equations
77629	12/18/2012	Crystal Sullivan	13615 Widover Ave.	Quinten, OK	73472	72A	Fuzzy Logic Concepts
77630	12/18/2022	Josh Blackburn	203 Park East Circle	Bedford, OH	44146	29A	Fundamental Problem Solving
77633	12/19/2012	Pathways Academy	2123 South Bender Ave.	Lafe, KY	42788	87A	Math Music
77634	12/19/2012	Pathways Academy	2123 South Bender Ave.	Lafe, KY	42788	11A	Making Cents of Money
77635	12/19/2012	Shae Brewer	6667 South Mendall	Carthage, MO	63197	57B	Linear Equations and Inequalities
77637	12/19/2012	Margit Burkhart	901 Metzler Ave.	New Haven, NJ	08912	33E	Audio Books
77638	12/20/2012	Justin Newberry	458 Court Lane	Indianapolis, IN	47963	78A	Math Music

Appendix E
One of Three Feedback Forms Used in the Optional Task

Feedback Form 1 – Task 1 - Rating the person who entered the Data

Rate the following statements	Strongly Disagree	Disagree	Agree	Strongly Agree	No Opinion
1. The person who entered this data was careful and precise.	❑	❑	❑	❑	❑
2. The invoices were bounded in the folder neatly.	❑	❑	❑	❑	❑
3. The Excel spread sheet was neat and easy to read.	❑	❑	❑	❑	❑
4. I had to fix very few errors.	❑	❑	❑	❑	❑
5. Most errors were small errors such as missing decimals.	❑	❑	❑	❑	❑
6. There were about the same amount of errors on every invoice entered.	❑	❑	❑	❑	❑
7. The amount of errors per invoice was about the same at the beginning of my work as the end of my work.	❑	❑	❑	❑	❑
8. I would recommend the person who entered the data for more work in this area.	❑	❑	❑	❑	❑

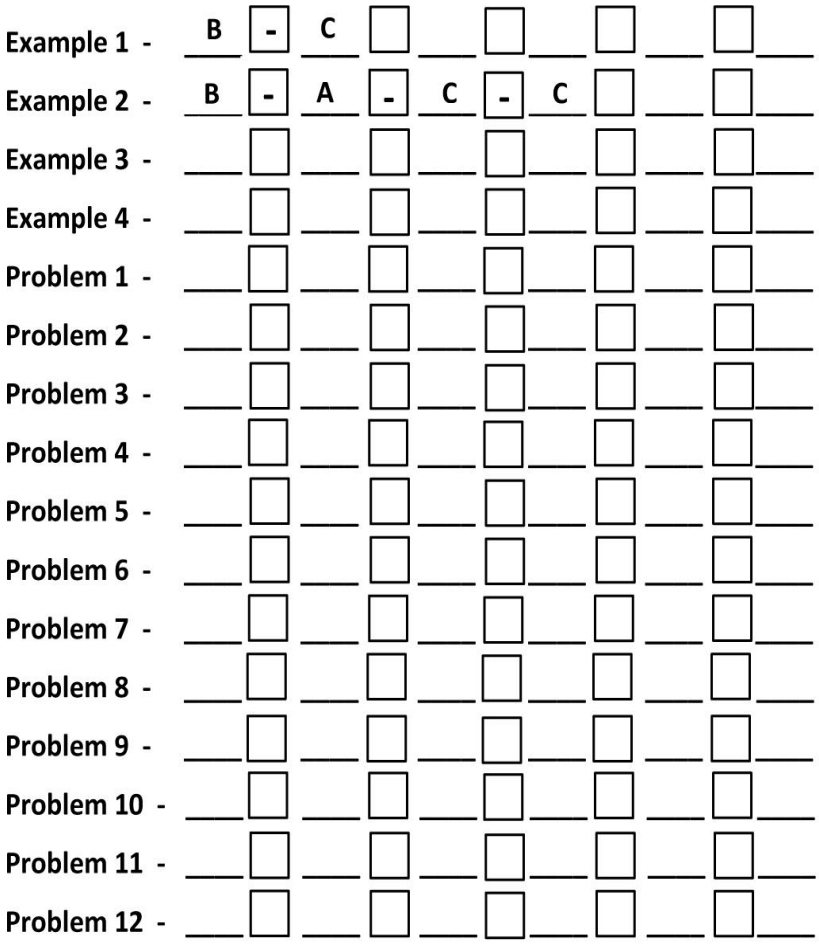
Appendix F
Water-jar Task Interface



Appendix G
Water-jar Task Progression Chart and Answer Sheet Form

Problems	Answers	Purpose	Notes
Example 1 & 2	Given to Participants	To introduce task	Answers pre-printed on answer form
Example 3 & 4	Participants Entered	To introduce task	Researcher checked before moving on
Problems 1-6	B-A-C-C	To test for pattern recognition	Hypothesis 1
Problems 7-9	B-A-C-C or either A-C or A+C	To test for Einstellung blindness	Hypothesis 2
Problems 10-12	All different and increasing in complexity	To test complex problem solving ability	Hypothesis 3

Water-Jar Answer Sheet Form



Appendix H
Matrix task for testing honesty

9.50	4.92	6.47	6.84	8.99	7.24	6.00	6.23	4.94	0.12	8.07	2.02
9.37	6.09	8.15	2.38	7.68	6.65	8.83	9.01	7.96	1.71	2.20	3.44
3.11	0.50	7.54	9.60	8.56	5.47	0.86	4.04	0.99	8.88	9.96	8.29
4.41	8.11	9.35	1.01	1.76	3.92	4.25	1.42	6.06	9.18	8.92	1.17
6.79	4.15	8.95	3.36	4.20	0.06	0.48	6.40	8.36	1.82	2.44	7.36
4.06	5.82	4.34	1.57	8.39	3.35	2.42	9.72	6.92	1.10	8.87	3.37
4.93	4.18	5.18	0.61	1.43	5.29	5.21	2.57	7.65	8.02	1.93	9.16
3.23	8.56	1.80	8.43	6.97	6.75	9.81	1.64	3.58	2.49	1.97	2.64
6.46	0.89	6.92	7.46	0.78	1.08	1.51	3.64	1.86	1.09	5.74	3.45
2.02	0.52	0.37	9.97	3.02	1.89	7.19	7.13	4.56	8.82	6.53	6.44
0.07	3.54	0.45	5.21	0.64	7.27	1.48	7.09	2.96	5.12	7.01	4.31
3.39	4.80	7.46	9.22	7.87	2.29	2.30	8.18	8.14	6.55	5.63	8.83
0.28	1.71	7.31	7.21	3.24	3.31	7.80	3.12	3.59	4.95	2.10	7.65
0.14	1.93	9.72	1.10	8.12	9.00	1.34	9.81	2.96	9.02	8.33	2.97
8.27	9.39	2.48	7.10	7.12	7.75	2.86	4.42	9.31	9.61	5.61	9.61
8.66	1.12	2.34	5.13	8.90	3.80	6.88	6.44	5.67	2.12	5.52	5.05
2.08	0.28	8.60	7.48	4.98	5.32	3.84	8.22	1.97	9.44	6.71	4.29
5.02	5.00	3.93	7.11	2.94	6.92	5.48	6.98	5.77	6.93	9.34	1.28
3.40	4.98	7.44	2.52	9.49	8.57	8.03	1.31	0.92	8.36	6.85	9.28
6.98	8.61	0.94	2.69	3.86	6.01	6.37	6.59	0.28	0.56	8.89	4.92