ESSAYS ON THE MOTIVATING EFFECTS OF GOALS IN ACCOUNTING

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Vikrant Vir Anand

January 2014

ESSAYS ON THE MOTIVATING EFFECTS OF GOALS IN ACCOUNTING

Vikrant Vir Anand, Ph. D.

Cornell University 2014

This dissertation presents two studies on the motivating effects of goals in accounting settings. Chapter 1 demonstrates that varying the horizon of a performance standard, which is treated as a goal by workers, affects motivation. In a lab experiment, participants exerted more effort on a task when facing short goal horizons. The shorter horizons allowed workers to achieve more goals in a given time, and this motivated them. However, as uncertainty in the task was increased, effort decreased with short goal horizons. By contrast, varying uncertainty did not affect effort when goal horizons were longer. The finding that uncertainty interacts with goal horizon can explain mixed findings in the goal-setting literature, and can possibly explain the variation in performance standard horizons observed in practice (De Angelis and Grinstein 2011). Chapter 2 demonstrates that the frequency at which goal progress information is provided to workers can also affect their motivation. In a lab experiment, motivation to achieve a goal decreased when feedback about goal progress was given frequently and there was high uncertainty in the effort-performance relationship. By contrast, motivation was not affected by uncertainty when feedback about goal progress was given less frequently because the uncertainty was able to cancel itself out to a greater extent. The reduction in effort due to frequent feedback and high uncertainty was due to participants' choice to abandon their goal and instead engage in leisure, especially when early realizations of uncertainty were negative and early feedback indicated low performance. These findings imply a benefit to aggregating information: when goal

progress information is aggregated over a longer horizon, some of the task uncertainty cancels itself out, and workers are less likely to give up on their goals. These two studies have implications for accounting research and accounting practice. Accountants need to carefully consider the frequency at which performance information is provided to workers, and also need to consider the amount of time given to workers to achieve their goals. Both studies suggest a benefit to aggregation of results when uncertainty is high, but a motivating effect of frequent goals when uncertainty is lower.

BIOGRAPHICAL SKETCH

Vikrant "Vic" Anand was born in San Antonio, Texas on May 8, 1974. He obtained a Bachelor of Science degree from the Massachusetts Institute of Technology in 1995. He received a Master of Science in Industrial Administration from Carnegie Mellon University's Graduate School of Industrial Administration (now the Tepper School of Business) in 2000, with university honors. He received the Vera Heinz Merit Scholarship while at Carnegie Mellon and was inducted into the Beta Gamma Sigma honor society. Vic began his doctoral studies at the Johnson Graduate School of Management at Cornell University in 2006 under Dr. Robert Bloomfield, Nicholas H. Noyes Professor of Management and Professor of Accounting. Vic accepted a tenure track position at Emory University's Goizueta Business School in Atlanta, Georgia in summer 2012. He currently teaches managerial accounting in Emory's BBA program.

I dedicate this dissertation to my loving wife Anne. I lovingly acknowledge her patience and support through the trials and tribulations of my doctoral program. I also dedicate this dissertation to Kenn Sperra, who has been like a second father to me. Or maybe more like a crazy uncle.

ACKNOWLEDGMENTS

I thank my dissertation committee members Rob Bloomfield (chair), Kevin Hallock, Mark Nelson, and Ted O'Donoghue for valuable guidance and advice. I also thank Ramji Balakrishnan, Ola Bengtsson, Emily Bianchi, Tim Brown, David De Angelis, George Easton, Kathryn Kadous, Eva Labro, Joe Pacelli, Harry Segal, Kristy Towry, Sean Wang, Ken Whelan, Melissa Williams, and seminar participants at the University of Alberta, The University of Iowa, Indiana University, Emory University, and Georgia Tech for helpful comments and suggestions. I thank Julie Mohamed for invaluable research assistance. I gratefully acknowledge financial support from Cornell University's Johnson Graduate School of Management and Emory University's Goizueta Business School.

TABLE OF CONTENTS

CHAPTER 1
PERFORMANCE STANDARD HORIZONS, UNCERTAINTY, AND CONTROLLABILITY

1-1. Abstract

The horizon of a performance standard specifies the time given to achieve the standard. Performance standard horizons vary considerably in practice, and the psychology literature provides mixed results on whether short or long horizons are more effective in motivating workers. In this paper I use an experiment with a real-effort task to show that the motivating effect of performance standards is determined by an interaction between the horizon of the standard and the amount of uncertainty in the relationship between effort and performance. When uncertainty was low, participants exerted significantly more effort when assigned performance standards with short horizons compared to long horizons. However, when uncertainty was high, the motivational effect of short horizons decreased because participants perceived themselves as having less ability to influence achievement of their standards through their efforts. The results demonstrate that the horizon of performance standards can affect worker motivation, and that horizons should be set as short as permitted by uncertainty in the situation. These findings provide an explanation for variation in performance standard horizons observed in practice, and suggest that investments in accounting systems that reduce noise in performance measures can allow management to set shorter horizon standards. The results also extend to the managerial accounting literature by demonstrating that the perceived controllability of a performance standard is affected by the matching of the standard's time horizon with the amount of noise in the worker's ability to achieve the standard. Finally, the results clarify mixed findings about horizon effects in the goal-setting literature.

1-2. Introduction

A performance standard is a threshold level of performance that must be achieved within a given period of time, known as the horizon. Considerable variation in the horizon of performance standards is observed in practice (De Angelis and Grinstein 2011) and the psychology literature provides mixed results on whether short or long horizons are more effective in motivating workers. This study proposes and tests the idea that the motivating effect of horizon varies with the level of uncertainty in workers' effort-performance relationship. Thus, while many previous studies have documented the effects that the *level* of a performance standard has on worker behavior (e.g. Anderson et al. 2010; Bol et al. 2010), this study demonstrates that the *horizon* of a performance standard can also have a measurable impact.

In this paper, I use an experiment in which participants exert cognitive effort to complete a task to show that the motivating effect of performance standards is determined by an interaction between the horizon of the standard and the amount of uncertainty in the relationship between effort and performance. When uncertainty was low, participants facing short horizons exerted significantly more effort than those facing long horizons, even though the rate of work demanded by the standards did not differ. However, when uncertainty was high, participants' effort did not vary with the horizon of their assigned standards. Debriefing responses indicate that participants felt they had little ability to control standard achievement through their efforts when facing short horizons and high uncertainty. However, uncertainty did not alter their perception of controllability when they faced long-horizon standards.

The results clarify the mixed findings in the goals literature on the effect of horizon. Some studies have found that goals with short horizons are more effective at motivating workers because frequent goal achievement provides satisfaction and increases task enjoyment, and because workers increase effort as they get closer to achieving a goal (see, e.g., Bandura and Schunk 1981; Heath et al. 1999). However, other studies have found that long-horizon goals lead to higher motivation and performance (e.g. Kirschenbaum et al. 1981) and conjecture that short-horizon goals reduce workers' flexibility to allocate effort over time. These authors argue that the loss of flexibility can lead workers to miss many of their short-horizon goals, which then decreases motivation. Thus, the literature suggests an interaction between goal horizon and uncertainty but has never tested it. This paper provides the first direct test of this idea and thereby contributes to the goal setting literature.

The results also contribute to a central topic in managerial accounting – controllability. The controllability principle states that a worker should be compensated based on that which he can control, and that failure to adhere to this principle can lead to dysfunctional behavior, such as reduced motivation (Merchant 1985). In this paper, I demonstrate that the perceived controllability of a performance standard is affected by the interaction of the standard's time horizon with the amount of noise in the worker's effort-performance relationship. When uncertainty was low, short-horizon standards were perceived as more controllable than long-horizon standards. However, when uncertainty was high, the reverse was true. The pattern of effort (and therefore motivation) exhibited by participants was identical to the pattern of perceived controllability. My findings are therefore consistent with the controllability principle.

In the experiment, 140 participants performed a computer-based, real-effort task of counting the number of zeros in a sequence of tables of 1's and 0's. They were awarded points for each table counted correctly and paid by the point at the end of the experiment (i.e. a piece rate). Participants were assigned performance standards that were displayed prominently while they worked but that had no effect on their monetary compensation. The experiment employed a 2×2 between-subjects design, varying standard horizon (short, long) and environmental uncertainty (high, low), and holding financial incentives constant in all experimental conditions. To vary standard horizon, half of the participants were assigned 8 standards and given 5 minutes to achieve each one (totaling 40 minutes). The other half were assigned 2 standards and given 20 minutes to achieve each one (also totaling 40 minutes). To vary environmental uncertainty, participants faced "shocks" that adjusted their point total every five minutes by a random number. For the participants assigned to the low-uncertainty condition, this random number was between -2 and 2; for the participants assigned to the high-uncertainty condition, the number was between -12 and 12. Thus, the shocks had a mean of zero for all, but higher variance for some.

Results indicate a significant interaction in which the effect of horizon on effort (measured as the number of tables counted correctly) depends on the level of uncertainty. Consistent with the hypothesis that short-horizon standards are more motivating when uncertainty is low, participants in the short-horizon condition counted significantly more tables than their counterparts in the long-horizon condition when uncertainty was low. Consistent with the hypothesis that the motivating effect of short-horizon standards is reduced when uncertainty is increased, participants in the short-horizon condition counted no more tables than their counterparts in the long-horizon condition when uncertainty was high. Debriefing responses indicated that uncertainty had a different impact on participants' perceptions of control over standard achievement. Specifically, uncertainty significantly reduced participants' perception of control when they faced short-horizon standards, but had no impact when they faced long-horizon standards. Debriefing responses also indicated that participants with short-horizon standards enjoyed the task significantly more and were more willing to participate again, consistent with the robust finding in the goals literature that goals increase task enjoyment and can make a boring task seem fun (Harackiewicz et al. 1984; Locke and Latham 2002).

This paper makes three contributions. First, this paper shows that the horizon of a standard can have an impact on motivation, even if the level of the standard is held constant. In practice, variation in the horizon of standards is observed and can range from a day, or even an hour, to multiple years. The findings in this paper demonstrate that the horizon of a standard must be matched to the level of uncertainty in the environment, with shorter horizons preferred whenever uncertainty allows it. To the extent that investments in information systems can reduce uncertainty in measuring worker performance, shorter-horizon standards can be employed.

Second, the paper shows that the actual and perceived controllability of a performance standard varies with the interaction between the standard's horizon and the uncertainty in the worker's production function. Holding environmental variables constant, a performance standard with a short horizon may be far less controllable than a performance standard with a long horizon. This lack of controllability has the potential to reduce worker motivation, which could in turn reduce the efficiency of the worker's incentive scheme.

Finally, the paper demonstrates that performance standards can affect worker motivation even when they are not tied directly to an extrinsic incentive such as a bonus. Workers internalize performance standards as goals and then derive satisfaction from achieving the goals. The satisfaction that workers derive from internal goal achievement could allow managers to supplement monetary pay with non-monetary utility.

The remainder of the paper is organized as follows. Section 1-3 discusses prior literature and develops hypotheses. Section 1-4 discusses my research design. Section 1-5 discusses my results, and Section 1-6 concludes.

1-3. Background and hypotheses

The U.S. Office of Personnel Management (OPM) defines a performance standard as "*a management-approved expression of the performance thresholds, requirements, or expectations that must be met to be appraised at a particular level of performance*" (U.S. Office of Personnel Management 1998). OPM further states that a key attribute of a performance standard is its timeliness (horizon), which specifies "*when or by what date the work is produced.*" In other words, a performance standard specifies a level of performance, but also requires specification of a time by which the level of performance must be met. Performance standards may be based on measures of inputs, such as hours worked, or outputs, such as profitability, production quantity or quality, uptime, and customer satisfaction. Managers use performance standards to communicate their desires and facilitate planning, to extract employees' private information, and for compensation, formal and informal performance evaluation, and promotion (Sprinkle 2003; Anderson et al. 2010; Murphy 2000; Indjejikian and Nanda 2002).

1-3.1 Performance standards and goals

Performance standards are distinct from the psychological construct of a "goal," but the two are closely linked. A worker's decision to attempt to achieve a performance standard represents an internalized desire, and this internalized desire meets Locke et al. (1981) definition of a goal: "*what an individual is trying to accomplish; it is the object or aim of an action.*" Locke (2001) argues that whenever a person makes a conscious choice to take an action that is directed towards an end, he has set a goal. Thus, the standard is an external stimulus, and the goal is the internal psychological state that results from it.

Architects of performance-standards systems appear to realize that standards create goals for workers. For example, Heckman et al. (1997) state that the primary function of performance-standards systems is to establish goals: "*By defining goals and providing incentives for achieving these goals, performance-standards systems attempt to bring to public-sector agencies the type of discipline that markets bring to firms.*" Sprinkle (2003) also acknowledges the goal-creating role of performance standards when he refers to the "*goals contained in accounting budgets and standards.*"

Goals are a critical part of the human cognitive subsystem through which people regulate their behavior (Bandura 1989). A large literature in psychology shows that achieving a goal is its own reward, providing non-monetary utility (Locke and Latham 1990; Locke and Latham 2002; Heath et al. 1999) and enhancing interest in the task (or equivalently, decreasing the cost of effort). Examples abound of cases where workers increased their performance simply because they were given performance standards. Latham (2004) reports that loggers who were assigned challenging standards cut down more trees and missed work less often than those told to "do your best" or those given no specific requirement. Kerr and Landauer (2004) describe General Electric's system of "stretch goals" (which would more-accurately be called "stretch standards") that are used as motivators but not always directly tied to pay. These anecdotes are representative of hundreds illustrating that performance standards can motivate workers even when not tied to extrinsic incentives. These studies also demonstrate that managers must carefully consider the motivational effects of performance standards, even if their own intent is to use the goal as a basis for communication, planning, or providing extrinsic incentives.

1-3.2 Choosing performance standard horizon

Managers must make a number of decisions when setting performance standards. One of the most studied is the *level* of the standard, which can range from a minimum performance standard to an easy-to-meet standard to a "stretch" standard (e.g. Merchant and Ferreira 1988; Anderson et al. 2010; Kerr and Landauer 2004). In this paper I focus on a far less-studied decision – the length of time that an employee is given to achieve the standard, called the *horizon*. In practice, there is considerable variation in the horizon of observed standards. For example, De Angelis and Grinstein (2011) analyze CEO compensation contracts and report that CEO performance standards range from a quarter to almost eight years.

1-3.2.1 Efficacy of short horizons

Since workers who commit to achieving performance standards set internal goals, I rely on the literature on goal setting to predict the effect of standard horizon on motivation. Numerous studies on goal horizon have been conducted over the last three decades yet the evidence on the motivating effect of goal horizon[1] remains mixed (Locke and Latham 1990). Some authors argue that since goals become more motivating when one is closer to achieving them, shorter goal horizons are more effective (e.g. Heath et al. 1999). Heath et al. (1999) propose that goals serve as reference points and that people experience sensations of gains or loss based on whether they exceed or fail to achieve their goals. Through numerous experiments, Heath et al. (1999) demonstrate that the gain/loss function of goals appears to be identical to the Prospect theory value function (Kahneman and Tversky 1979). The Prospect theory value function is convex in the loss domain and concave in the gain domain, and therefore exhibits the property of diminishing sensitivity. In the loss domain, where the value function is convex, the closer one gets to the reference point, the higher is the marginal utility. If goals function as Heath et al. (1999) propose, then the closer one is to achieving a goal, the more it will motivate them.

Other studies also argue that decomposing a goal into multiple subgoals means that individuals can achieve more goals and therefore attain more satisfaction from goal achievement (Locke and Latham 2002). Bandura and Simon (1977) argue that people intuitively realize the efficacy of short-horizon goals, because a number of experimental studies report that subjects assigned long-horizon goals set short-horizon goals for themselves.

Finally, another line of studies argues that short horizons prevent procrastination and thereby increase effort (e.g. Bandura 1989). Many recent papers in behavioral economics suggest that individuals set personal goals as commitment devices to help mitigate the effects of present bias, or self-control problems (e.g. Jain 2009; Hsiaw 2010; Koch and Nafziger 2009; Suvorov and van de Ven 2008). This stream is part of a growing literature suggesting that management control systems not only align incentives of managers and workers, but also help workers with their own self-control problems (Kaur et al. 2010a).

[1] The literature on goal setting has adopted the terms "proximal" and "distal" to refer to short-horizon and long-horizon goals, respectively. In the more recent behavioral economics literature, authors refer to the process of setting goal horizon as "goal bracketing." An overarching goal is "bracketed" into multiple subgoals. In this paper, I use the term horizon to refer to the length of time available to achieve a goal.

1-3.2.2 Inefficacy of short horizons

While the above literature provides evidence that short horizons are often effective in motivating workers and improving performance, other studies argue that long-horizon goals are preferable in at least some circumstances. One line of research argues that longer horizon goals afford more flexibility in allocating effort over time (Kirschenbaum 1985; Read et al. 1999). For example, Kirschenbaum et al. (1981) conducted an experiment in which they assign study goals to undergraduates. They found that those assigned monthly study goals spent more time studying and improved grades more than those assigned daily goals or no goals. They conjecture but do not test the idea that daily goals were incompatible with the irregular schedules of their subjects, and that after failing to achieve initial daily goals, the subjects gave up.

Read et al. (1999) survey the literature on how to frame choices and reach a similar conjecture. They argue that when longer horizons are considered, people are better able to make tradeoffs across choices and that this leads to higher returns to effort. For example, they state that "*artists and academics have days when the muse is resident, and days when she is painfully absent. In such situations, the most efficient way for workers to organize their time is to work long hours when the return to time spent is high, and take leisure when the return is low*." Consider what would happen if the artist/academic in this hypothetical scenario had daily output goals. On days when the muse was absent, the artist/academic would fail to achieve his goal, and on other days, he would far exceed his goal. If the artist/academic was loss-averse, the sensations of loss experienced on days when he failed to achieve his goals would outweigh the sensations of gain on the days when he exceeded his goals and he would have been better off with a longer goal horizon. Under longer horizons, random variation in outcomes can cancel out through the power of large sample sizes, thereby giving people more control over goal achievement through their efforts.

1-3.3 A theory of optimal performance standard horizons

The choice of optimal performance standard horizon is based on the theory that performance standards lead workers to set internal goals, and that their responses to the properties of those goals will determine the observed responses to the standard. Therefore, in the following paragraphs, I assume that workers internalize the goal of achieving their assigned performance standards and discuss the effects of horizon and uncertainty on the motivating effect of those goals. I also assume that achieving the performance standard has no implications for compensation, performance evaluation or other extrinsic incentives.

The evidence for and against short goal horizons implies a tradeoff between the greater motivation of having nearby goals and the loss of flexibility that accompanies short goal horizons. Goals with shorter horizons are more motivating when uncertainty in the worker's effort-output relationship is low. However, when uncertainty is high, workers do not have time to respond to uncertainty in the environment and therefore cannot achieve short-horizon goals. I therefore propose that the motivating effect of performance standards is determined by an interaction between horizon and uncertainty because that interaction determines the controllability of the standard. This conjecture relies on the assumption that workers decide to achieve the performance standard assigned to them and therefore set a goal based on it.

I derive predictions about the interaction between horizon and uncertainty by first considering a traditional economic framework and then incorporating one straightforward behavioral assumption: that people derive utility from achieving a goal. In a traditional economic framework, a rational agent maximizes expected utility, which derives from changes in wealth. A performance standard that is not tied to wealth would have no effect on agent effort. However, I assume that people will internalize performance standards as goals and derives utility from goal achievement. There is support for this assumption in the literature. For example, several behavioral economic models explicitly assume a fixed increase in utility from goal attainment (e.g. Jain 2009), and in their review of goal setting theory, Locke and Latham (2002) explicitly state that "*the more goal successes one has, the higher one's total satisfaction.*" My assumption is much simpler than that used in other recent behavioral economic models of goals. For example, Hsiaw (2010) and Koch and Nafziger (2009) both follow the Heath et al. (1999) model of goals as reference points and assume that people experience a sensation of gain or loss that is increasing in the amount by which they exceed or fall short of their goal. In contrast, I derive my hypotheses simply by assuming that simply achieving a goal results in a "burst" of utility that is greater than when the goal is missed.

If achieving more goals leads to higher utility, then in the absence of uncertainty, shorter goal horizons, which mean more goals, increase returns to effort. As a result, people will work harder to achieve a series of short-horizon goals than they would to achieve a single long-horizon goal that is equivalent to the sum of the short horizon goals. This leads to the first hypothesis:

> **Hypothesis 1**: Shorter goal horizons increase effort when uncertainty about the effort-output relationship is low.

Hypothesis 1 suggests that performance standard horizons should be as small as possible when uncertainty is low. However, uncertainty in the effort-output relationship imposes a lower bound on the optimal horizon of the performance standard. In a world with no uncertainty, workers know exactly what the results of their efforts will be and can decide whether the benefits of achieving a performance standard exceed the costs. When uncertainty is present, the relationship between effort and output becomes stochastic. A worker who exerts high effort might not achieve his performance standard, and as performance standard horizons are shortened, the control the worker has over achieving his standard through his efforts is diminished. The worker simply may not have enough time to exert the effort required to guarantee that he achieves his performance standard. By contrast, if the horizon is longer, the worker has the flexibility to respond to the realizations of uncertainty that he observes. It is also possible that uncertainty may cancel itself out over a longer horizon.

Consider a worker with a convex cost of effort who is paid a piece rate for his output. When there is no uncertainty in the environment, the relationship between effort and output is deterministic. The worker will choose a level of effort at which the marginal benefit (the piece rate) just equals the marginal cost of effort. Call this level of effort e^*. Since achieving a goal provides extra utility (by assumption), the worker can be induced to work a little harder if given a goal. He will work until the marginal benefit (piece rate plus goal utility) just equals the marginal cost. Call this level of effort e_g^*. If uncertainty is introduced into the effort-output relationship, where uncertainty has a mean of zero, a rational worker will reduce his effort. The reason is that, because of the uncertainty, he is no longer guaranteed to achieve the goal if he exerts e_g^*. At e_g^*, the probability of achieving the goal and realizing goal utility is less than 1. Thus, the expected utility from goal achievement is reduced, and he will exert effort in between e^* and e_g^*.

To see the differential effect on workers with short- and long-horizon goals, consider a setting where a worker faces two consecutive periods, each with a goal of g, and uncertainty that realizes at the end of each period. By the reasoning above, the worker will reduce effort as uncertainty (the variance in the distribution) is increased. Now consider the effect on a worker who faces a single goal of $2g$ for both periods. The worker can respond to the uncertainty realized after the first period. If the uncertainty is negative, he can increase effort. If it is positive, he can decrease effort. A rational worker will realize that the sum of the two draws might cancel each other out. His control over the attainment of his goal of $2g$ is therefore less affected by the uncertainty than his counterpart who faced two goals of g, and his reduction in effort in response to an increase in uncertainty will be less.

Thus, increasing uncertainty reduces the controllability of performance standards for workers facing short-horizon standards, but has a smaller effect on controllability for workers facing long-horizon standards. The reduction in controllability will reduce motivation (Bandura and Wood 1989). This leads to hypothesis 2.

Hypothesis 2: Greater uncertainty in the effort-output relationship reduces the motivational advantage of shorter goal horizons.

In this paper, I do not posit a main effect of uncertainty on effort. A series of analytical papers in neoclassical economics considered the problem of labor supply under uncertainty, but without the presence of a goal (e.g. Block and Heineke 1973; Tressler and Menezes 1980). One prediction of these papers is that when uncertainty is additive (e.g. a fixed salary plus a variable bonus), risk-averse workers increase effort to hedge against low realizations of uncertainty (i.e. an income effect). Some related empirical papers support this finding (e.g. Parker et al. 2005). However, because none of these papers considered the effects of goals, I do not rely on them for my predictions. If an income effect arises because of uncertainty in this experiment, it would weaken the directional effects hypothesized above, and work against finding a result.

1-4. Method

1-4.1 Participants

Participants were 140 students at a large, private research university. Participants were recruited through the web site of the business school's laboratory, which maintains a large pool of subjects for use in experimental research. Since the task required no specific skills or knowledge beyond basic computer skills, the ability to exert effort on a straightforward task, and the ability to respond to a typical pay-for-performance arrangement, this subject pool was an appropriate choice (Libby et al. 2002). Upon seeing advertisements for studies at the lab's web site, members of the subject pool voluntarily sign up for experiments. This study's advertisement stated "*In this experiment you will be asked to perform a computer-based task for money. We expect that the study will take 75 minutes. This study is only for pay (no credit). Expect to earn about $16.*"[2]

[2] A supplementary analysis in section 1-5.5.3 examines whether participants had a reference point of $16 when they arrived in the lab. It concludes that many did come with a goal of $16, but that the likelihood of arriving with this goal did not differ across treatments.

1-4.2 Task overview and experimental design

The goal of the experiment was to examine the impact of performance standards on real effort, and how that impact varies with uncertainty and standard horizon. To elicit real effort, participants were asked to perform a straightforward, repetitive task for which they received monetary compensation[3]. Participants were awarded points for each repetition of the task completed, and gained or lost points as a result of periodic random events. Points were redeemed for cash at the end of the experiment using a fixed, known exchange rate. Participants were also given performance standards but these were not tied to their pay – there was no bonus for achieving a standard or penalty for failing to do so.

Participants' effortful task was to count the number of zeros in tables comprised of 54 randomly generated zeros and ones (see Figure 1-1). The task is adapted from Abeler et al. (2011), who argued that it was useful for assessing real effort because it "*does not require any prior knowledge, performance is easily measurable, and there is little learning possibility; at the same time, the task is boring and pointless, and we can thus be confident that the task entailed a positive cost of effort for subjects.*" The positive cost of effort is desirable for this study since the theory employs an assumption common to most agency models – that effort is costly (Lambert 2001).

[3] In a real-effort task, participants must exert cognitive or physical effort in order to accomplish the task. For example, Kachelmeier et al. (2008) required participants to design rebus puzzles, a cognitively demanding task. Ariely et al. (2008) had participants assemble structures out of Legos, a task that required physical and cognitive effort. The alternative to a real effort task is one where the experimenters proxy for effort by asking participants to choose a number. For example, Anderhub et al. (2002) study reciprocity in a principal-agent setting; participants assigned to the role of agents chose a number between 0 and 20 to represent their effort choice. Effort in their experiment was costly in that it reduced agents' payout, but the different levels of effort did not require different levels of cognitive exertion. This practice of choosing a number is quite common in experimental economics (see, e.g. Fehr et al. 1997; Frederickson 1992).

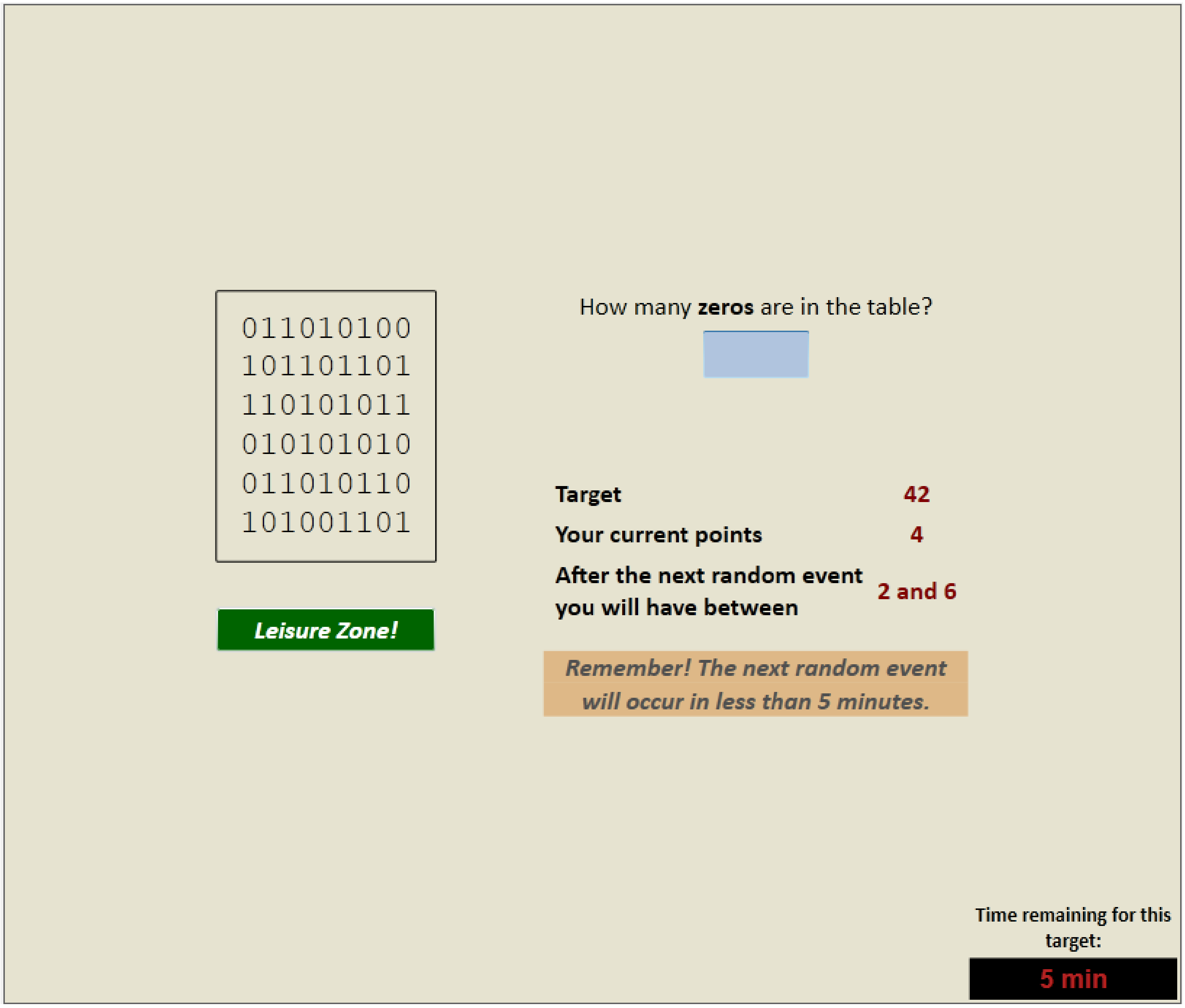


Figure 1-1: The screen on which participants performed their task during the main round of the experiment. The table is the box with 1's and 0's. Participants had to count the number of zeros in the table and enter it into the textbox labeled "How many zeros are in the table?" If they entered a correct answer, the word *Correct* appeared under the textbox, the table flashed, and a new table was displayed. If they entered an incorrect answer, the word *Wrong* appeared along with "Tries left: x", where x was 2 or 1 (a total of 3 tries were allowed for each table). The timer in the lower right corner of the screen updated every minute, as did the box with "Remember!" The screen for the preliminary round differed in that there was no *Leisure Zone* button, no reminder of an upcoming random event, no timer in the bottom right hand corner of the screen, and the only feedback text was "Your current points".

The experiment used a 2 × 2 between-subjects design. Performance standard horizon was manipulated at two levels, short and long. Participants in the *short-horizon* condition were given 8 standards to achieve and 5 minutes to achieve each one. Participants in the *long-horizon* condition were given 2 standards to achieve and 20 minutes to achieve each one. Uncertainty was manipulated at two levels, low and high. After every 5 minutes, a random number was added to each participant's point total. For participants in the *low-uncertainty* condition, this random number was drawn from a discrete uniform distribution with support [-2, 2]. For participants in the *high-uncertainty* condition, this random number was drawn from a discrete uniform distribution with support [-12, 12]. To ensure that realizations of random events were the same across the four cells of the design, each cell used the same 35 sequences of random numbers (a different one for each participant in the cell)[4]. The dependent variable, effort, was measured by the number of tables counted correctly in the main round. A timeline of the experiment in the short- and long-horizon conditions is shown in Figure 1-2.

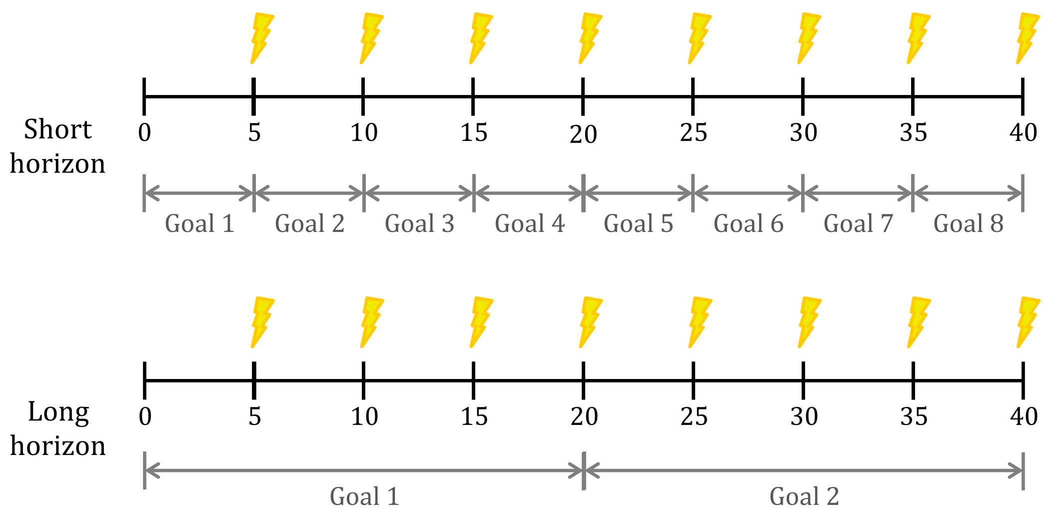


Figure 1-2: Timeline of the task. Participants counted tables of 1's and 0's for 40 minutes. Each table was worth 2 points. At the end of every 5-minute period, participants experienced a "shock" – a random number was added to their points; this is represented by a lightning bolt symbol in the timelines above. Shocks in the *low (high) uncertainty* condition were drawn from discrete U[-2, 2] (U[-12, 12]). In the *short-horizon* conditions, participants were given a performance standard (or goal) of earning 42 points every 5 minutes. In the *long-horizon* conditions, participants were given a standard of earning 168 points every 20 minutes. Participants in the *long-horizon* conditions experienced 4 shocks per goal, whereas those in the *short-horizon* conditions experienced 1 shock per goal and had no opportunity to react to the shock before goal evaluation.

1-4.3 Task and procedures

Upon entering the lab, each participant was seated at a partitioned computer station, assigned to a treatment using a predetermined random-assignment scheme, and verbally

[4] Each sequence was drawn from a continuous, uniform distribution with support [-1, 1], and then scaled up to either [-2, 2] or [-12, 12], depending on the uncertainty level for that participant. The number was also rounded down to the nearest integer.

instructed to "*follow the instructions on the screen*." The experimental session was divided into two rounds: a 10-minute preliminary round and a 40-minute main round. The preliminary round required participants to count tables without any performance standards or random events. The main round was identical to the preliminary round, except that it included the performance standards and random events described above. The preliminary round reduces noise in the analysis of effort by giving participants an opportunity to learn the task before the main round, and by providing a baseline measure of individual variation in performance before performance standards are introduced.

1-4.3.1 Preliminary round

Instructions for the preliminary round (see appendix in section 1-8) described the rules governing the task and incentives. The instructions informed participants that, after counting the zeros in a table, they should enter their count into a textbox (see screen snapshot in Figure 1-1). If the participant entered the correct count, the computer displayed the next table and incremented the participant's point total by 2 points. If the participant entered an incorrect count, he was given two more tries to count the same table before a penalty of 2 points was assessed. The penalty scheme, adopted from Abeler et al. (2011), discourages effort avoidance through guessing and skipping tables that look difficult to count. Participants were also informed that they would be paid a show-up fee of $10 plus an additional 15¢ for every 10 points earned.

After the preliminary round ended, the software informed participants of the number of tables they had counted correctly and their pay for the preliminary round. They then progressed to the main round of the experiment.

1-4.3.2 Main round

After completing the preliminary round, participants were shown instructions for the main round, which are reproduced in the appendix in section 1-9. These instructions informed participants that they would continue to perform the same task of counting tables of 1's and 0's, that each table would continue to be worth 2 points, that a penalty of 2 points would be assessed after 3 incorrect answers, and that the pay rate would remain at 15¢ for every 10 points received. The instructions then informed participants of changes from the preliminary round.

1-4.3.2.1 Performance standards

All participants were informed that they would now be given performance standards to achieve[5]. Participants in the *short-horizon* conditions were told that their standard was to earn 42 points every 5 minutes for the next 40 minutes (8 standards). Participants in the *long-horizon* conditions were told that their standard was to earn 168 points every 20 minutes for the next 40 minutes (2 standards)[6]. They were told that each standard was independent of the others – points earned on an early standard would not carry over to future standards. Finally, *participants were informed that there would be no bonus for achieving a standard or penalty for failing to achieve a standard. It was emphasized that pay would be determined solely by total points accumulated*.

1-4.3.2.2 Random events

Participants were then informed that since "*unexpected events can affect work performance … a random number will be added to – or subtracted from – your point total after every 5 minutes.*" Participants in the *low-uncertainty* condition were informed that this random number would be between -2 and 2, and those in the *high-uncertainty* condition were informed that the random number would be between -12 and 12. These point additions were referred to as "random events."

---

[5] Instructions to subjects described performance standards as targets. See page 2 of the main round instructions in section 1-9.

[6] To determine the level of the performance standards, I conducted a calibration test, which was an extended preliminary round. A total of 21 participants counted tables for 20 minutes. Since counting times improved during the test, I analyzed data from the last 5 minutes of the test to control for learning. I found that participants required 14.5 seconds on average to count a table of ones and zeros. A participant who counts at that rate during the main round count slightly more than 20 tables in 5 minutes (1 table / 15 sec x 60 sec / min x 5 min). Since two points were awarded per table, I reasoned that participants should be able to earn at least 40 points in 5 minutes, and therefore set the performance standard for 5 minutes at 42 points.

When choosing levels of uncertainty, my goal was to have a large enough difference in uncertainty between the short-horizon conditions that I would be able to detect an effect, but not so large that participants in the long-horizon, high-uncertainty condition would perceive no control over target achievement. To estimate the effect of uncertainty under different goal horizons, I performed simulations in *Mathematica*. In the short-horizon condition, participants facing low uncertainty who counted 21 tables (and therefore earned 42 points) had a 60% chance of achieving their target. By counting one additional table, they improved their chance of target achievement by 40% and guaranteed target achievement. By contrast, participants in the short-horizon, high uncertainty condition who counted 21 tables had 52% chance of target achievement and each additional table counted increased the chance of target achievement by only 8%. Thus, if controllability is thought of as the marginal effect of effort on likelihood of target achievement, participants in the short-horizon, low-uncertainty condition had 5 times greater control over target achievement through effort than their counterparts in the short-horizon, high-uncertainty condition (40% versus 8%). By design, controllability in the long-horizon conditions varied less with uncertainty; participants in the long-horizon, low-uncertainty condition had 2 times greater control over target achievement than their counterparts in the long-horizon, high-uncertainty condition (44% versus 22%).

The mean of small samples of random numbers often deviates significantly from the population mean, and such deviation occurred for many participants. Twenty-five percent of participants (35 / 140) experienced a significant net loss to their points when their random numbers turned out to be mostly negative. Approximately 30% (41 / 140) experienced a significant net addition to their points. The remaining participants experienced a mix of positive and negative additions to their points that summed to approximately zero. The distribution of the sum of random events for all participants is shown in Figure 1-3.

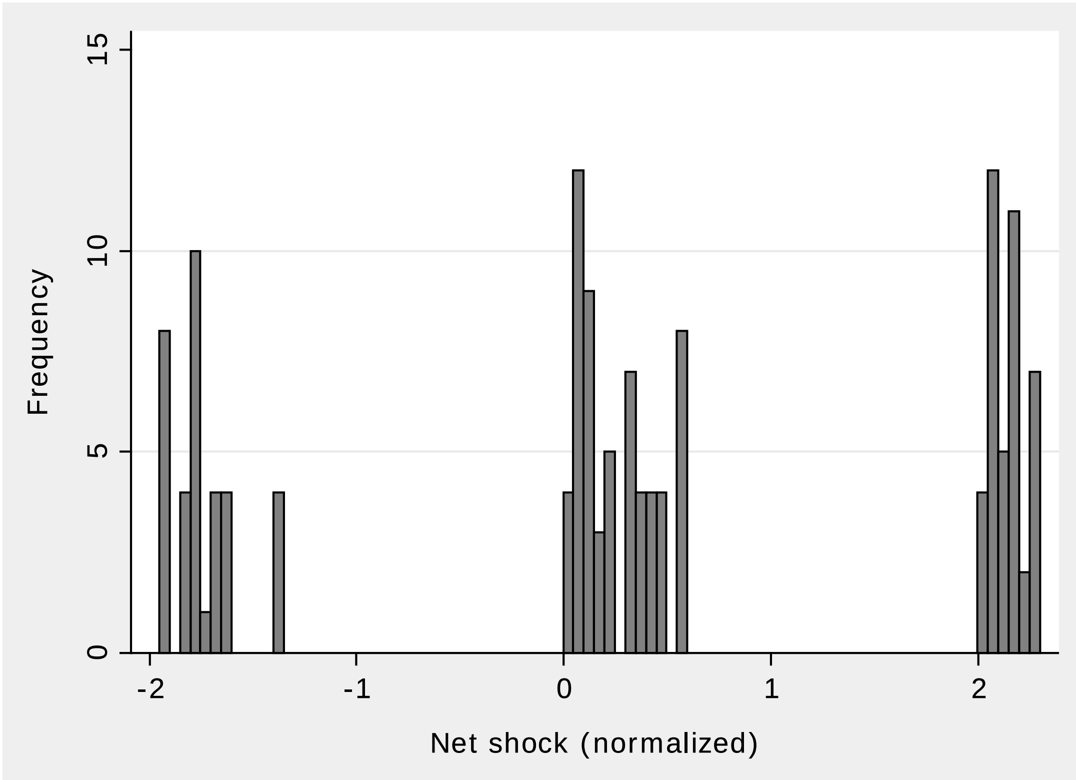


Figure 1-3: Distribution of the net, or total, shock experienced by participants. During the main round, a random number was added to each participant's point total after every 5 minutes. In the *low (high) uncertainty* condition, this random number was uniformly distributed between -2 and 2 (-12 and 12). Each participant experienced 8 random changes ("shocks") to their points, and this graph shows the distribution of the sum of these shocks across all participants.

All shocks were initially drawn from U[-1, 1] and then scaled to [-2, 2] or [-12, 12] depending on the experimental condition. This figure shows a histogram of the sum of the normalized shocks (i.e. the sum of the U[-1, 1] numbers) experienced by participants. In the low uncertainty condition, the left (right) cluster corresponds to a net point addition of approximately -4 (+5), and the middle cluster corresponds to a net point addition of zero. In the high uncertainty condition, the left (right) cluster corresponds to a net point addition of approximately -20 (+26), and the middle cluster corresponds to a net point addition of approximately +4.

Four participants were excluded from this graph because they had extreme realizations of net shock (normalized net shock of -5.8).

1-4.3.2.3 Real-time feedback

While counting tables, participants were given real-time feedback about progress towards their current performance standard. As depicted in the screen snapshot in Figure 1-1, on-screen feedback displayed the current standard, the current point level, and the distribution of points after the next random event. For example, if a participant in the *low-uncertainty* condition had 4 points, the screen reminded him that after the next random event, he would have between 2 and 6 points.

Finally, the screen displayed time remaining to achieve the current standard, and the timer was updated every minute.

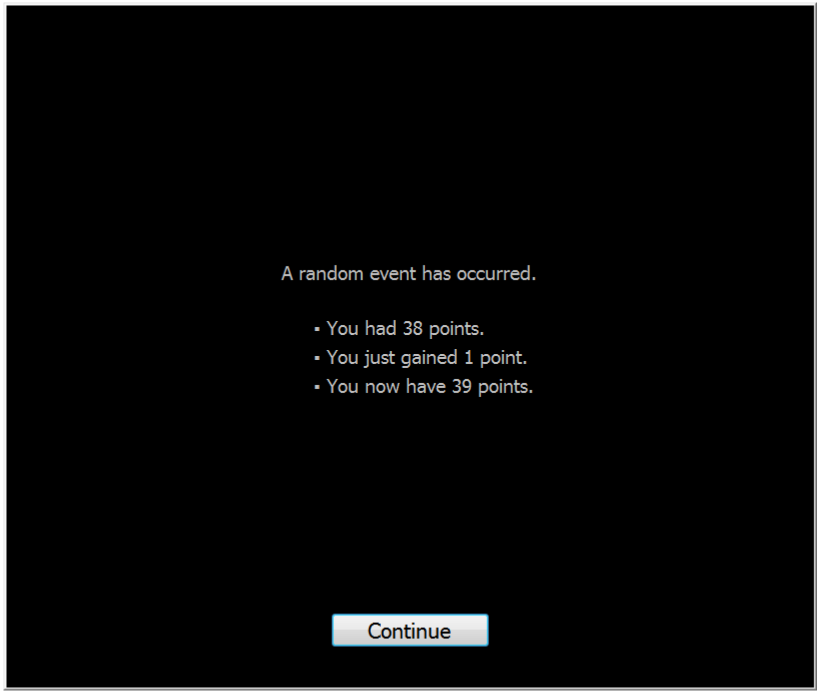


Figure 1-4: The Notification window appeared every 5 minutes during the 40-minute main round. This window informed participants of the value of the random event, or "shock" that occurred, as well as their new point total.

1-4.3.2.4 Notification of random events and standard achievement

After every 5 minutes of the main round, a window appeared notifying participants of the random event that occurred; it showed the amount of the point gain or loss and the new point total (see Figure 1-4). If time to achieve a standard had elapsed, this window also told participants whether or not they had achieved their standard. Standard achievement was announced in bold, green letters, and animated fireworks were displayed on the screen (see Figure 1-5). Failure to achieve a performance standard was announced in dark red letters with no fireworks (see Figure 1-6). A timeline of the main round that indicates the relative timing of notifications of random events and standard achievement is shown in Figure 1-2.

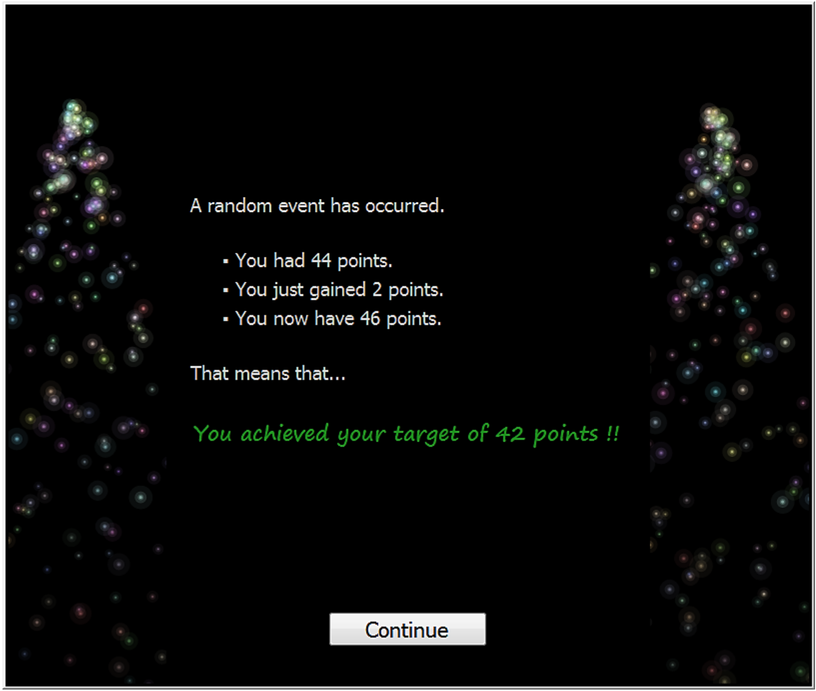


Figure 1-5: The Notification window appeared every 5 minutes during the 40-minute main round. This window informed participants of the value of the random event, or "shock" that occurred, as well as their new point total. When the time for a goal had elapsed, this window also informed participants of whether or not they achieved their goal. This example shows performance standard achievement. When standard achievement occurred, animated fireworks were displayed on the screen for 15 seconds.

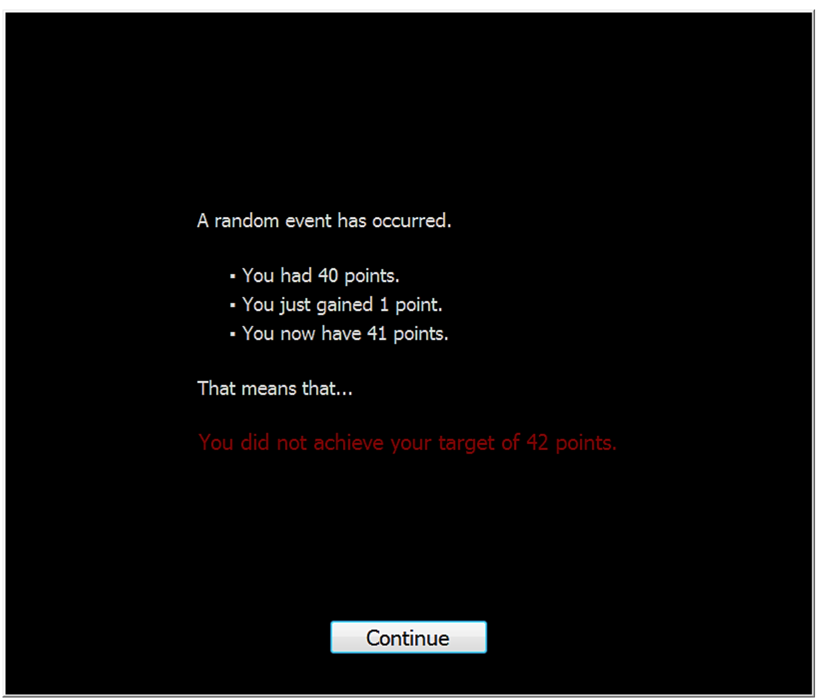


Figure 1-6: The Notification window appeared every 5 minutes during the 40-minute main round. This window informed participants of the value of the random event, or "shock" that occurred, as well as their new point total. When the time for a performance standard ("target") had elapsed, this window also informed participants of whether or not they achieved their standard. This example shows a failure to achieve the standard.

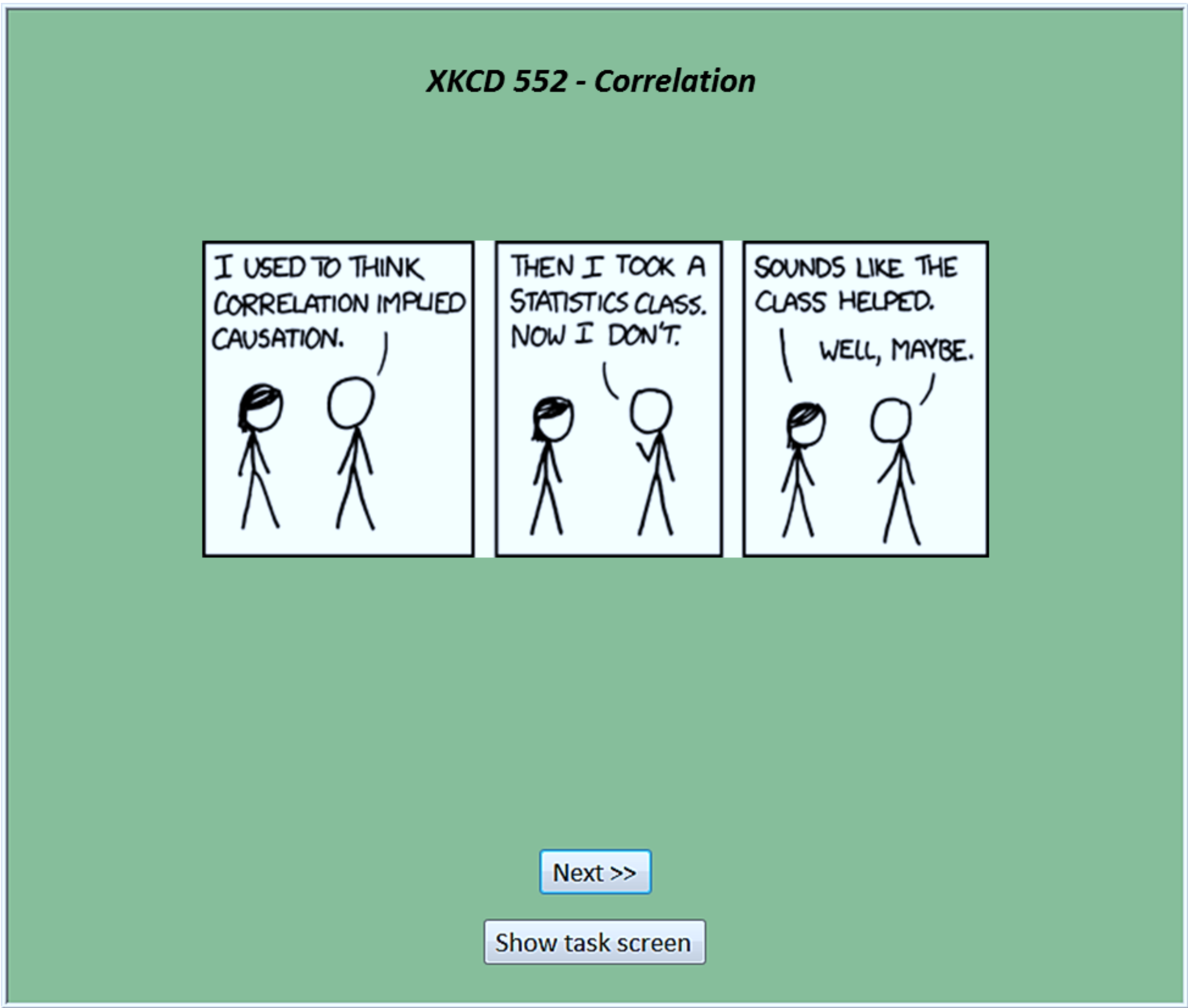


Figure 1-7: Leisure Zone screen displaying a comic with statistical humor. This screen completely covered the main work screen (Figure 1-1) so participants could not earn points while in the Leisure Zone. The sequence of content appeared random from participants' perspective, but was predetermined and was the same for all participants. Once participants clicked the Next button, they could not return to the previous content. Participants could resume work at any time by clicking the Show task screen button.

1-4.3.2.5 The Leisure Zone

Participants were given the choice to either count tables or engage in an alternative task that I named the *Leisure Zone*. The purpose of the *Leisure Zone* was to obtain a clear measure of time spent working. For example, if participants wished to temporarily stop working to reward themselves for achieving a standard or to procrastinate on a current standard, they could enter the *Leisure Zone*. Participants were told that they could take a break from counting tables at any time during the main round by clicking on a button titled "*Leisure Zone*" on the main screen shown in

Figure 1-1. Clicking the button opened a window that displayed a sequence of content from popular websites (see Figure 1-7). The content consisted of comics from two strips (XKCD and Calvin and Hobbes), celebrity gossip from TMZ.com, whimsical photos of animals, recent news stories, and pictures from a fashion website (The Sartorialist). To entice participants to enter the *Leisure Zone*, they were shown sample content while reading the instructions. They were told that content would appear in random order[7] and that they could stay in the *Leisure Zone* as long as they liked. There was no penalty for time spent in the *Leisure Zone*, but it was not possible to earn points while in there.

1-4.3.2.6 Debriefing and final payment

After completing the 40-minute main round, participants were shown a screen that informed them of the total points they had accumulated from counting tables and from random events. The screen informed them of their total pay, which was equal to the $10 show-up fee plus the total of accumulated points multiplied by the point-cent exchange rate. Participants then answered 14 debriefing questions. After completing these questions, the software instructed the participant to go to the lab's office, where the experimenter paid the participant in cash[8]. The participant then left the lab.

1-4.3.3 Comprehension and manipulation checks

Participants were asked nine questions to assess their understanding of the main round. These questions asked participants about their assigned standards (level of each standard, number of standards, time allotted for each standard, independence of each standard from the others), determinants of pay, random events (distribution, frequency, possible impact on standard achievement), and the *Leisure Zone* (when it would be available and the effect on their pay). On

[7] The actual sequence of content was predetermined and was the same for all participants.

[8] After each participant completed debriefing questions, the software sent the experimenter an email notifying him of the participant's computer number and final pay. This allowed the experimenter to remain in the office in case many participants finished at the same time and had to line up to be paid.

every question, more than 85% of participants provided a correct answer on the first try; on most questions only one or two participants out of 140 provided an incorrect answer on the first try. The responses to the comprehension check questions indicate that participants understood the instructions. Most importantly, participants understood that their pay was determined only by the total points they accumulated and not by standard achievement. Ninety-four percent of participants (131/140) answered this question correctly on the first try. Identical inferences can be drawn from the results reported below if these nine participants are excluded.

1-4.4 Software

The task, instructions, comprehension check questions, and debriefing questions were administered using custom software written in the C# programming language. The software was designed to allow participants to perform the experiment with minimal involvement from the experimenter. The software captured every keystroke and mouse movement made by each participant, and disabled the Windows system keys to prevent participants from using their assigned computer for any purpose but the experiment.

Table 1-1: Effect of performance standard horizon and uncertainty on effort

Participants counted tables of 1's and 0's in two rounds, a preliminary round and a main round. *Raw performance* refers to the number of tables counted correctly during the 40-minute main round. *Baseline performance* refers to the number of tables counted correctly during the last 4 minutes of the 10-minute preliminary round, multiplied by 10; this represents the number of tables the participant would have counted correctly in the main round if he had continued working with the same effort intensity as during the last 4 minutes of the preliminary round. *Baseline-adjusted performance* equals *raw performance* minus *baseline performance*. Each participant earned two points for every correctly counted table. During the main round, a random number was added to each participant's point total after every 5 minutes. In the *low (high) uncertainty* condition, this random number was uniformly distributed between -2 and 2 (-12 and 12). Participants were also given performance standards that did not affect their pay. *Short (long) horizon* refers to the experimental condition in which participants were given eight 5-minute (two 20-minute) standards of earning 42 (168) points.

Panel A shows performance in each of the experimental conditions. One, two, or three asterisks (*) next to a baseline-adjusted cell mean indicate that the mean is different from 0 at the 10%, 5%, or 1% level, respectively (two-tailed *t*-test). Panel B shows an ANOVA of *baseline-adjusted performance* on *horizon* and *uncertainty*. Panel C shows the results of the planned contrasts that were used to test hypotheses 1 and 2. Panel D shows an ANCOVA of *raw performance* on *horizon* and *uncertainty*, with *raw performance* as a covariate. Panel E shows tests of hypotheses 1 and 2 based on the ANCOVA in panel D.

Panel A: Cell means

	Raw performance				Baseline-adjusted Performance			
Goal horizon	N	Low uncertainty	High uncertainty	Diff	N	Low uncertainty	High uncertainty	Diff
Short	69	**173**	**168**	5	69	**45*** **	**29*** **	16
Long	71	**155**	**163**	-8	71	**18*** **	**29*** **	-11
Difference		18	5			27	0	

Panel B: ANOVA of baseline-adjusted performance on horizon and uncertainty

Num. of obs	140	R^2	0.0665
Root MSE	36.3568	Adjusted R^2	0.0459

Source	Partial SS	df	MS	F	Prob > F
Model	12,807.86	3	4,269.29	3.23	0.0245
Horizon	5,903.08	1	5,903.08	4.47	0.0364
Uncertainty	238.41	1	238.41	0.18	0.6717
Horizon x Uncertainty	6,520.27	1	6,520.27	4.93	0.0280
Residual	179,767.32	136	1,321.82		
Total	192,575.17	139	1,385.43		

Panel C: Tests of hypotheses using planned contrasts on baseline-adjusted performance

	Effect size	F	p-value (1-tailed)
Hypothesis 1:			
Short-horizon/low-uncertainty performance > long-horizon/low-uncertainty performance	26.64	9.53	0.0013
Hypothesis 2:			
(Short-horizon minus long-horizon performance at low uncertainty) > (Short-horizon minus long-horizon performance at high uncertainty)	27.30	4.93	0.0140

Panel D: ANCOVA of raw performance on horizon and uncertainty using baseline-adjusted performance as a covariate

Num. of obs	140	R^2	0.4720
Root MSE	34.6949	Adjusted R^2	0.4564

Source	Partial SS	df	MS	F	Prob > F
Model	145,276.69	4	36,319.17	30.17	0.0000
Baseline performance	138,671.60	1	138,671.60	115.20	0.0000
Horizon	5,619.29	1	5,619.29	4.67	0.0325
Uncertainty	87.59	1	87.59	0.07	0.7878
Horizon x Uncertainty	4,883.38	1	4,883.38	4.06	0.0460
Residual	162,504.20	135	1,203.73		
Total	307,780.89	139	2,214.25		

Panel E: Tests of hypotheses using planned contrasts on covariate-adjusted raw performance

	Effect size	F	p-value (1-tailed)
Hypothesis 1:			
Short-horizon/low-uncertainty performance > long-horizon/low-uncertainty performance	24.53	8.83	0.0018
Hypothesis 2:			
(Short-horizon minus long-horizon performance at low uncertainty) > (Short-horizon minus long-horizon performance at high uncertainty)	23.71	4.06	0.0230

1-5. Results

1-5.1 Cell means of dependent variables

Panel A of Table 1-1 reports cell means for raw performance, which is defined as the number of tables counted correctly by each participant during the main round of the experiment. The panel also reports means for *baseline-adjusted performance*, which is defined as *raw performance* less ten times the number of tables counted by each participant during the last four minutes of the preliminary round[9]. *Baseline-adjusted performance* reduces the substantial between-subject variation in productivity by accounting for the number of tables the participant would have counted in the 40-minute main round had he continued to work with the same effort intensity as in the preliminary round. Figure 1-8 presents *raw* and *baseline-adjusted performance* across the four cells of the design.

Mean *baseline-adjusted performance* is different from zero in every cell at the 1% level, indicating that participants' rate of counting tables in the main round was higher than that in the preliminary round. Table 1-1, panel B reports the results of an ANOVA for *baseline-adjusted performance*. *Horizon* is significant (F = 4.47, p = 0.0364), indicating that participants worked harder when provided with short-horizon goals. The interaction between *Horizon* and *Uncertainty* is also significant (F = 4.93, p = 0.0280), indicating that the effect of *Horizon* changes at different levels of *Uncertainty*.

1-5.2 Hypotheses tests

Table 1-1, panel C reports the results of the hypothesis tests. Hypothesis 1 states that when uncertainty is low, short-horizon performance will exceed long-horizon performance. I test hypothesis 1 with a planned contrast that $\mu_{ShortLow} > \mu_{LongLow}$, where μ_{XY} represents the mean of

[9] The choice of 4 minutes reflects a tradeoff. Performance continues to improve throughout the preliminary round, but shorter windows provide noisier measures of preliminary-round performance since participants were interrupted while counting their last table. Identical inferences can be drawn from the results reported below if the last 5, 3, or 2 minutes are used as the basis for measuring baseline performance.

baseline-adjusted performance with horizon of X and uncertainty of Y. Participants in the short-horizon/low-uncertainty condition did indeed work harder than their counterparts in the long-horizon/low-uncertainty condition. Their adjusted performance was higher by 26.64 tables ($F = 9.53$, $p = 0.0013$, one-tailed[10]). This result strongly supports hypothesis 1 and suggests that short-horizon goals are motivating when there is low task uncertainty.

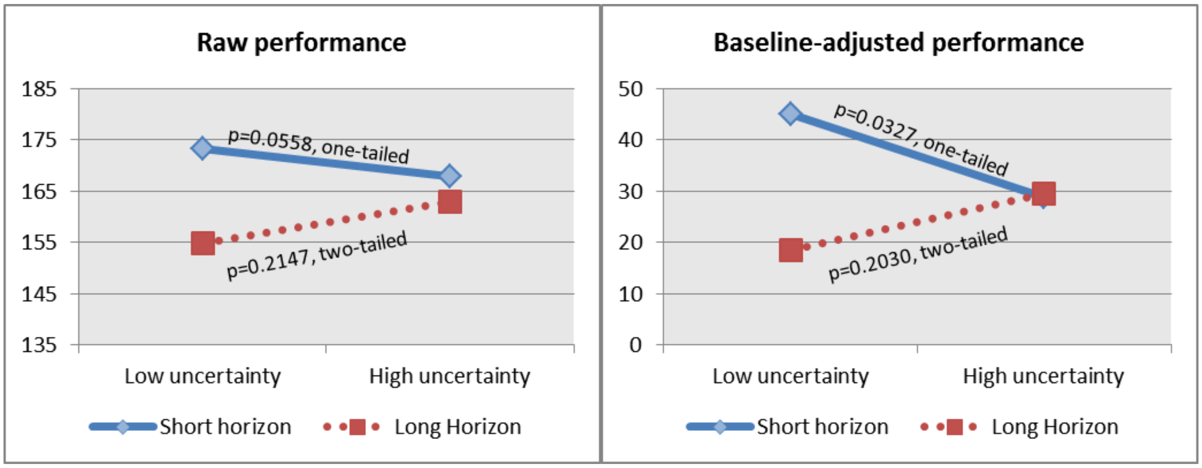


Figure 1-8: Participants counted tables of 1's and 0's in two rounds, a preliminary round and a main round. *Raw performance* refers to the number of tables counted correctly during the 40-minute main round. *Baseline performance* refers to the number of tables counted correctly during the last 4 minutes of the 10-minute preliminary round, multiplied by 10; this represents the number of tables the participant would have counted correctly in the main round had he continued to work with the same effort intensity as during the last 4 minutes of the preliminary round. *Baseline-adjusted performance* equals *raw performance* minus *baseline performance*. Each participant earned two points for every correctly counted table. During the main round, a random number was added to each participant's point total after every 5 minutes. In the *low (high) uncertainty* condition, this random number was uniformly distributed between -2 and 2 (-12 and 12). Participants were also given performance standards that did not affect their pay. *Short (long) horizon* refers to the experimental condition in which participants were given eight 5-minute (two 20-minute) standards of earning 42 (168) points.

The left graph shows the mean of *raw performance* in each experimental condition. The right graph shows the mean of *baseline-adjusted performance* in each experimental condition. P-values are simple effects of uncertainty at each level of horizon (i.e. whether the slope of the line differs from zero).

Hypothesis 2 predicts that the beneficial effect of short horizon standards will diminish as task uncertainty is increased. This hypothesis was tested by using a planned contrast of

[10] MacDonald (2011) states that the *F*-distribution with 1 numerator degree of freedom and *d* denominator degrees of freedom is equivalent to the square of the Student *t*-distribution with *d* degrees of freedom. Therefore, "*[a]s long as the F test has 1 numerator degree of freedom, the square root of the F statistic is the absolute value of the t statistic for the one-sided test.*" Since all contrast tests reported in this paper are Wald tests with 1 numerator degree of freedom, this method is appropriate, and is equivalent to dividing the two-tailed *p*-value in half.

$(\mu_{ShortLow} - \mu_{LongLow}) > (\mu_{ShortHigh} - \mu_{LongHigh})$. This hypothesis is also supported strongly. The performance benefit of short-horizon goals is an additional 27.3 tables when uncertainty is low compared to when it is high (F = 4.93, p = 0.0140, one-tailed). This result suggests that uncertainty provides a bound on the usefulness of shorter performance standard horizons. Support for the hypothesis is driven primarily because uncertainty reduces performance when horizons are short (untabulated but see Figure 1-8: F = 3.45, p = 0.0327, one-tailed). Performance under long-horizon goals increases as uncertainty increases, but the increase is not statistically significant (untabulated but see Figure 1-8: F = 1.64, p = 0.2030, two-tailed)[11].

1-5.2.1 Robustness of hypotheses

An alternative to analyzing *baseline-adjusted performance* is to use an ANCOVA in which *raw performance* is the dependent variable and *baseline performance* is a covariate to account for between-subject variation in performance when goals are absent. Table 1-1, panel D reports the results of this ANCOVA. There is a main effect of *horizon* (F = 4.67, p = 0.0325) and a statistically significant interaction between *horizon* and *uncertainty* (F = 4.06, p = 0.0460). Panel E shows that hypothesis 1 is supported and reveals a similar effect size as with ANOVA: performance in the short-horizon/low-uncertainty cell was higher than that in the long-horizon/low-uncertainty cell by 24.53 tables (F = 8.83, p = 0.0018, one-tailed), compared to 26.64 with ANOVA. Panel E shows that hypothesis 2 is also supported and reveals a similar effect size: the benefit of short-horizon standards over long-horizon standards decreases by 23.71 tables as uncertainty increases (F = 4.06, p = 0.0230, one-tailed), compared to 27.30 with ANOVA. Thus both hypotheses are supported and the effect sizes are similar to their counterparts in Table 1-1, panel C. I conclude that the use of

[11] I use a one-tailed test for the simple effect of uncertainty when horizon is short because the theory suggests that short-horizon goals become less effective when uncertainty increases. Since I offer no prediction for long-horizon goals, a two-tailed test is more appropriate for them.

baseline-adjusted performance as my dependent variable yields identical inferences and is easier to interpret[12].

Table 1-2: Perceived controllability of performance standards

After completing the main round of the experiment, participants were asked a series of debriefing questions. This table shows the response to the question on whether participants believed they had control over achievement of their assigned performance standards ("targets"). Participants were asked, "*How strongly do you agree with the following statement: 'I had control over whether I could achieve the targets.'*" Responses were solicited on a 7-point Likert scale in which -3 indicated strong disagreement, 0 indicated a neutral response, and 3 indicated strong agreement.

Panel A shows the mean response to the question on perceived controllability for the entire sample and for each tercile of net shock. An asterisk (*) next to a cell mean indicates that the mean is different from 0 at the 10% level or better. Panel B displays an ANOVA of the response to the question using *horizon* and *uncertainty* to explain the variation; it also shows planned contrasts.

Eight random numbers ("shocks") were added to each participant's points total after every 5 minutes of the main round. In the *low (high) uncertainty* condition, this random number was uniformly distributed between -2 and 2 (-12 and 12). Some participants experienced mostly negative shocks, others experienced mostly positive shocks, and others received a mix. Each participant's cumulative, or net, shock was computed and these were ranked into terciles. The distribution of net shocks is shown in Figure 1-3. Four subjects were dropped from this analysis because they experienced extreme values of net shocks (-13 in low uncertainty, and -72 in high uncertainty). Panel C displays a separate ANOVA and planned contrasts for each net shock tercile.

Panel A: Cell means of perceived controllability

	Entire sample				Negative net shocks			
Goal horizon	N	Low uncertainty	High uncertainty	Diff	N	Low uncertainty	High uncertainty	Diff
Short	67	**1.53***	**0.42**	1.11	16	**1.75***	**0.50**	1.25
Long	69	**1.40***	**0.26**	1.14	19	**1.90***	**-1.33***	3.23
Difference		0.13	0.16			-0.15	1.83	

	Neutral net shocks				Positive net shocks			
Goal horizon	N	Low uncertainty	High uncertainty	Diff	N	Low uncertainty	High uncertainty	Diff
Short	31	**1.87***	**0.25**	1.62	20	**0.91**	**0.67**	0.24
Long	29	**1.36***	**1.07***	0.29	21	**1.00***	**0.50**	0.50
Difference		0.51	-0.82			-0.09	0.17	

[12] In an untabulated ANCOVA, I test for, but do not find, a 3-way interaction between horizon, uncertainty, and baseline performance. Thus, variations in baseline performance do not have different effects in the different experimental conditions.

Panel B: ANOVA and contrasts of perceived controllability for entire sample

Num. of obs	136	R^2	0.0925
Root MSE	1.79551	Adjusted R^2	0.0719

Source	Partial SS	df	MS	F	Prob > F
Model	43.39	3	14.46	4.49	0.0049
Horizon	0.71	1	0.71	0.22	0.6398
Uncertainty	42.65	1	42.65	13.23	0.0004
Horizon x Uncertainty	0.01	1	0.01	0.00	0.9611
Residual	425.55	132	3.22		
Total	468.94	135	3.47		

	Effect size	F	p-value (1-tailed)
Contrast 1:			
(Short-horizon minus long-horizon response at low uncertainty) > (Short-horizon minus long-horizon response at high uncertainty)	-0.03	0.00	0.4806

Panel C: ANOVA and contrasts of perceived controllability by net shock tercile

ANOVA and contrasts of perceived controllability – Neutral net shocks

Num. of obs	60	R^2	0.1132
Root MSE	1.72627	Adjusted R^2	0.0657

Source	Partial SS	df	MS	F	Prob > F
Model	21.30	3	7.10	2.38	0.0790
Horizon	0.35	1	0.35	0.12	0.7320
Uncertainty	13.61	1	13.61	4.57	0.0370
Horizon x Uncertainty	6.58	1	6.58	2.21	0.1429
Residual	166.88	56	2.98		
Total	188.18	59	3.19		

	Effect size	F	p-value (1-tailed)
Contrast 1:			
(Short-horizon minus long-horizon response at low uncertainty) > (Short-horizon minus long-horizon response at high uncertainty)	1.33	2.21	0.0715

ANOVA and contrasts of perceived controllability – Negative net shocks

Num. of obs	35	R^2	0.4244
Root MSE	1.63036	Adjusted R^2	0.3686

Source	Partial SS	df	MS	F	Prob > F
Model	60.74	3	20.25	7.62	0.0006
Horizon	6.15	1	6.15	2.31	0.1385
Uncertainty	43.59	1	43.59	16.40	0.0003
Horizon x Uncertainty	8.53	1	8.53	3.21	0.0830
Residual	82.40	31	2.66		
Total	143.14	34	4.21		

	Effect size	F	p-value (1-tailed)
Contrast 1:			
(Short-horizon minus long-horizon response at low uncertainty) > (Short-horizon minus long-horizon response at high uncertainty)	-1.98	3.21	0.0415

ANOVA and contrasts of perceived controllability – Positive net shocks

Num. of obs	41	R^2	0.0121
Root MSE	1.88457	Adjusted R^2	-0.068

Source	Partial SS	df	MS	F	Prob > F
Model	1.62	3	0.54	0.15	0.9280
Horizon	0.01	1	0.01	0.00	0.9492
Uncertainty	1.40	1	1.40	0.39	0.5336
Horizon x Uncertainty	0.17	1	0.17	0.05	0.8286
Residual	131.41	37	3.55		
Total	133.02	40	3.33		

	Effect size	F	p-value (1-tailed)
Contrast 1:			
(Short-horizon minus long-horizon response at low uncertainty) > (Short-horizon minus long-horizon response at high uncertainty)	-0.26	0.05	0.4143

1-5.3 Perceived controllability of standards

The theory underlying hypotheses 1 and 2 suggests that the controllability of a standard is determined by the interaction between standard horizon and uncertainty in the environment. Uncertainty reduces the controllability of short-horizon standards by much more than it reduces that of long-horizon standards, because when the horizon is long, there is time to respond to uncertainty as it is realized.

I test this notion of controllability by analyzing the responses to a debriefing question on perceived controllability of performance standards. Participants were asked to rate agreement with the statement "*I had control over whether I could achieve the targets*" on a 7-point Likert scale that ranged from strong disagreement to strong agreement. Because some participants experienced mostly negative or mostly positive realizations of uncertainty (see Figure 1-3), I group the realizations of net (total) realizations into terciles and analyze the responses to this question separately for each tercile, as well as for the entire sample.

Table 1-2, panel A shows the mean response to this question for the entire sample and for each tercile of net shocks. Results from the entire sample show that participants in the low-uncertainty conditions clearly perceived the task as more controllable than those in the high-uncertainty conditions. The ANOVA in Table 1-2, panel B shows a significant main effect of uncertainty ($F = 13.23$, $p = 0.0004$), and the top-left graph of Figure 1-9 shows this pattern graphically. The means of the low-uncertainty cells are positive and significantly different from zero, indicating that these participants felt they had control over standard achievement. By contrast, the means of the high-uncertainty cells are not significantly different from zero, indicating a neutral perception of controllability.

The ANOVA for the entire sample (Table 1-2, panel B) does not show an interaction between uncertainty and horizon ($F = 0.00$, $p = 0.9611$), due largely to the great variation in perceptions of controllability when shocks consistently hinder or aid standard attainment. To get a better look at perceived controllability in the absence of such effects, I examine the tercile of participants who experience neutral net shocks. The top-right graph of Figure 1-9 shows that perceived controllability in the short-horizon, low-uncertainty cell is significantly higher than that in the short-horizon, high-uncertainty cell (untabulated but see Figure 1-9: $F = 6.79$, $p=0.0059$, one-tailed)). However, when standards have long horizons, there is no simple effect of uncertainty (untabulated but see Figure 1-9: $F = 0.21$, $p=0.6524$, two-tailed). A planned contrast (Table 1-2, panel C, neutral net shocks) shows that short-horizon standards are perceived as more controllable than long-horizon standards when uncertainty is low, but that the difference in perceived controllability declines as uncertainty is increased ($F = 2.21$, $p = 0.0715$, one-tailed). Thus, consistent with the theory underlying hypotheses 1 and 2, controllability of short-horizon standards is strongly affected by uncertainty, but controllability of long-horizon standards is not affected.

When uncertainty consistently hinders standard achievement (negative net shocks), participants in the high-uncertainty conditions unsurprisingly report that the standards are less controllable. Table 2, panel C, negative net shocks, shows a significant main effect of uncertainty ($F = 16.40$, $p = 0.0003$). Surprisingly, participants in the long-horizon/high-uncertainty condition found the standards the least controllable. The bottom-left graph of Figure 1-9 shows the pattern, and an ANOVA in Table 1-2, panel C, negative net shocks, confirms a significant interaction between horizon and uncertainty when in the most negative tercile of net shock ($F = 3.21$, $p = 0.0830$). Furthermore, the mean response for the long-horizon/high-uncertainty condition is -1.33 (Table 1-2, panel A) and this is significantly different from zero ($t = -2.53$, $p = 0.0353$, two-tailed). I conjecture that the effect of uncertainty on perceived controllability is strongest among participants with long-horizon standards because they received consistently negative shocks that repeatedly pulled them back from each standard, making it difficult for them to achieve their goals.

By contrast, those with short horizon standards were able to achieve many of their goals because some shocks were positive, even if the sum of all their shocks was negative.

Participants who received consistently positive shocks report that controllability is the same, regardless of experimental condition. The ANOVA in Table 1-2, panel C, positive net shocks, shows no significant main effects of horizon and uncertainty, and no interaction. The bottom-right graph of Figure 1-9 shows the pattern of responses graphically. The responses for three of the four cells do not differ significantly from zero. This suggests that participants recognized that uncertainty weakened the link between effort and standard achievement, even though the uncertainty was helpful.

1-5.4 Task enjoyment and willingness to participate again

The theory underlying hypotheses 1 and 2 assumes that people enjoy having goals to work towards, and that achieving more goals leads to more satisfaction (i.e. utility) than achieving fewer goals. I address this aspect of the theory by analyzing data from two debriefing questions. The first question asked participants to rate their agreement with the statement: "*I enjoyed my task of counting tables and trying to achieve targets.*" The second asked them to rate their agreement with the statement: "*I would participate in this experiment again.*" Both questions were answered on a 7-point Likert scale that ranged from strong disagreement (-3) to strong agreement (+3), with zero indicating a neutral response. Given that the task was the same for all participants, differences in responses between experimental conditions must be due to the effects of horizon and uncertainty.

Table 1-3, panel A shows the mean responses to the question on task enjoyment by experimental condition, and Figure 1-10 displays the pattern of responses. Consistent with the theory, there is a significant main effect of horizon ($F = 4.09$; $p = 0.0452$); participants in the short-horizon condition reported significantly higher enjoyment of the task than participants in the long-horizon condition. In fact, the responses of participants in the short-horizon condition do not differ significantly from zero (untabulated: $\mu = 0.21$, $t = 0.5756$, $p = 0.2844$), indicating neither enjoyment nor dislike of the task, while participants in the long-horizon condition gave responses that were

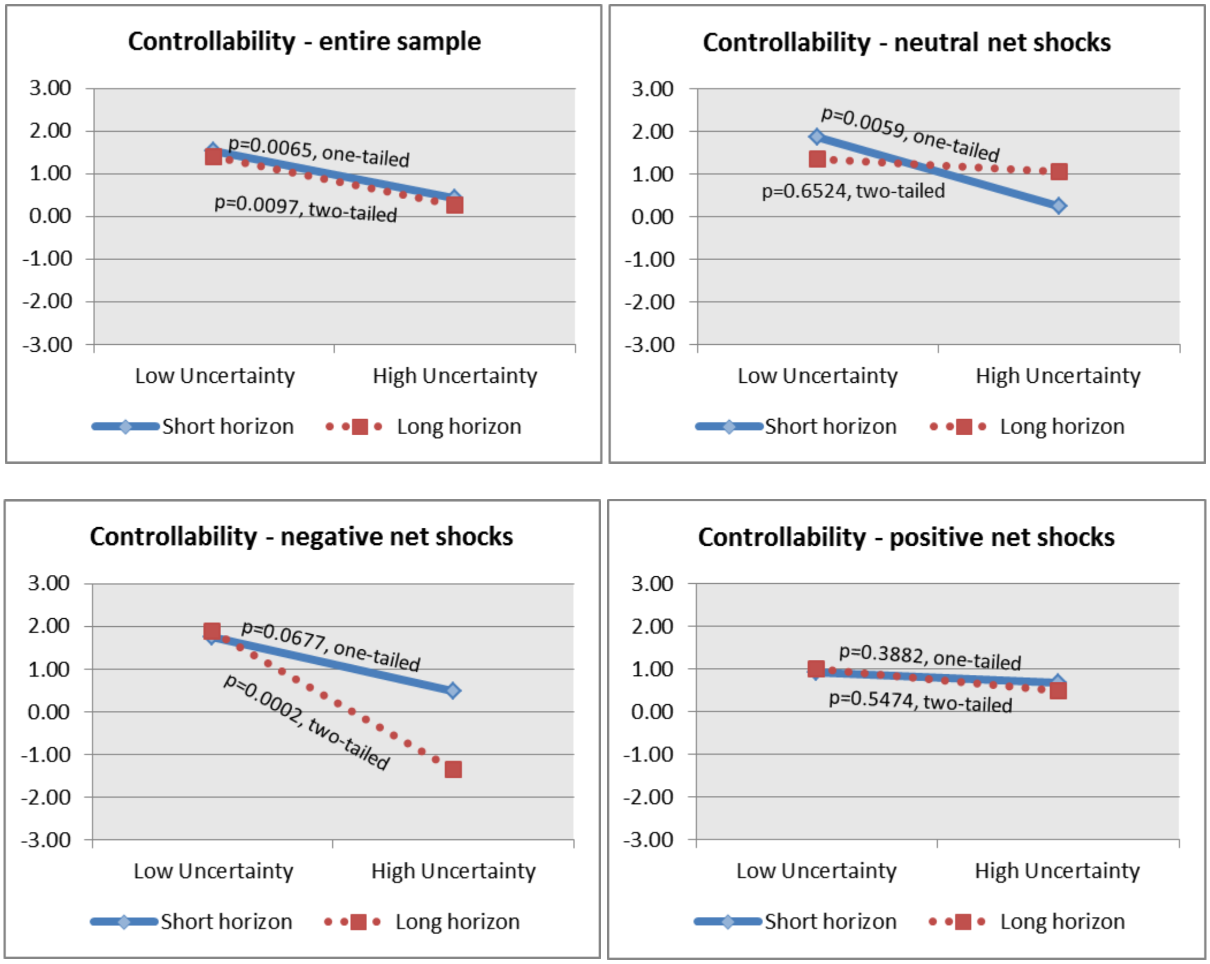


Figure 1-9: Participants counted tables of 1's and 0's in two rounds, a preliminary round and a main round. Participants earned two points for every correctly counted table. During the main round, a random number was added to each participant's point total after every 5 minutes. In the *low (high) uncertainty* condition, this random number was uniformly distributed between -2 and 2 (-12 and 12). Participants were also given performance standards that did not affect their pay. *Short (long) horizon* refers to the experimental condition in which participants were given eight 5-minute (two 20-minute) standards of earning 42 (168) points.

After completing the main round of the experiment, participants were asked a series of debriefing questions. The figures show the response to the question on whether they believed they had control over performance standard achievement. Participants were asked, "*How strongly do you agree with the following statement: 'I had control over whether I could achieve the targets.'*" Responses were solicited on a 7-point Likert scale in which -3 indicated strong disagreement, 0 indicated a neutral response, and 3 indicated strong agreement.

The top left graph shows the mean response by experimental condition using the entire sample. Going clockwise, the other graphs are for participants who experienced neutral, positive, and negative net shocks, respectively. P-values are simple effects of uncertainty at each level of horizon (i.e. whether the slope of the line differs from zero).

negative and significantly different from zero (untabulated: $\mu = -0.80$, $t = -2.5026$, $p = 0.0173$), indicating dislike of the task. The first contrast test in Table 1-3, panel A shows that participants in the short-horizon/low-uncertainty cell enjoyed the task significantly more than participants in the long-horizon/low-uncertainty cell ($F = 4.40$, $p = 0.0189$, one-tailed), providing support for the theory that more goals provide more utility, and that goals increase task enjoyment.

Table 1-3: Task enjoyment and willingness to participate again

After completing the main round of the experiment, participants were asked a series of debriefing questions. Panel A shows the response to the question on task enjoyment. Participants were asked, "*How strongly do you agree with the following statement: 'I enjoyed my task of counting tables and trying to achieve targets.'*" Panel B shows the response to the question on participants' willingness to participate again. Participants were asked, "*How strongly do you agree with the following statement: 'I would participate in this experiment again.'*" Responses to both questions were solicited on a 7-point Likert scale in which -3 indicated strong disagreement, 0 indicated a neutral response, and 3 indicated strong agreement.

The first table in each panel shows the mean response in each experimental condition. An asterisk (*) next to a cell mean indicates that the mean is different from 0 at the 10% level or better. The second table in each panel shows an ANOVA of responses on *horizon* and *uncertainty*. The third table in each panel shows the results of two planned contrasts on the responses, based on the preceding ANOVA.

Panel A: Analysis of debriefing question on task enjoyment

Cell means of task enjoyment

Goal horizon	N	Low uncertainty	High uncertainty	Diff
Short	67	**0.21**	**-0.24**	0.45
Long	69	**-0.76***	**-0.62***	-0.14
Difference		0.97	0.38	

ANOVA and contrasts of task enjoyment

Num. of obs	136	R^2	0.0375
Root MSE	1.99098	Adjusted R^2	0.0156

Source	Partial SS	df	MS	F	Prob > F
Model	20.37	3	6.79	1.71	0.1675
Horizon	16.21	1	16.21	4.09	0.0452
Uncertainty	0.60	1	0.60	0.15	0.6976
Horizon x Uncertainty	3.38	1	3.38	0.85	0.3575
Residual	523.25	132	3.96		
Total	543.62	135	4.03		

	Effect size	F	p-value (1-tailed)
Contrast 1:			
Short-horizon/low-uncertainty response > long-horizon/low-uncertainty response	1.01	4.40	0.0189
Contrast 2:			
(Short-horizon minus long-horizon response at low uncertainty) > (Short-horizon minus long-horizon response at high uncertainty)	0.63	0.85	0.1788

Panel B: Analysis of debriefing question on willingness to participate again

Cell means of willingness to participate again

Goal horizon	N	Low uncertainty	High uncertainty	Diff
Short	67	**1.18***	**1.27***	-0.10
Long	69	**0.43**	**0.53**	-0.10
Difference		0.75	0.74	

ANOVA and contrasts of willingness to participate again

Num. of obs	136	R^2	0.0359
Root MSE	1.97815	Adjusted R^2	0.014

Source	Partial SS	df	MS	F	Prob > F
Model	19.23	3	6.41	1.64	0.1837
Horizon	18.89	1	18.89	4.83	0.0297
Uncertainty	0.33	1	0.33	0.08	0.7719
Horizon x Uncertainty	0.00	1	0.00	0.00	0.9946
Residual	516.53	132	3.91		
Total	535.76	135	3.97		

	Effect size	F	p-value (1-tailed)
Contrast 1:			
Short-horizon/low-uncertainty response > long-horizon/low-uncertainty response	0.75	2.47	0.0594
Contrast 2:			
(Short-horizon minus long-horizon response at low uncertainty) > (Short-horizon minus long-horizon response at high uncertainty)	0.00	0.00	0.4973

Figure 1-10: Participants counted tables of 1's and 0's in two rounds, a preliminary round and a main round. Participants earned two points for every correctly counted table. During the main round, a random number was added to each participant's point total after every 5 minutes. In the *low (high) uncertainty* condition, this random number was uniformly distributed between -2 and 2 (-12 and 12). Participants were also given performance standards that did not affect their pay. *Short (long) horizon* refers to the experimental condition in which participants were given eight 5-minute (two 20-minute) standards of earning 42 (168) points.

After completing the main round of the experiment, participants were asked a series of debriefing questions. The left figure shows the mean responses by experimental condition to the question on task enjoyment. Participants were asked, "*How strongly do you agree with the following statement: 'I enjoyed my task of counting tables and trying to achieve targets.'*" The right figure shows the mean responses by experimental condition to the question on participants' willingness to participate again. Participants were asked, "*How strongly do you agree with the following statement: 'I would participate in this experiment again.'*" Responses to both questions were solicited on a 7-point Likert scale in which -3 indicated strong disagreement, 0 indicated a neutral response, and 3 indicated strong agreement. P-values are simple effects of uncertainty at each level of horizon (i.e. whether the slope of the line differs from zero).

Table 1-3, panel B shows the mean responses to the question on willingness to participate again by experimental condition, and Figure 1-10 displays the pattern of responses. The pattern of results is identical to that for the question on task enjoyment. There is a main effect of uncertainty (F = 4.83; p = 0.0297), with participants in the short-horizon condition indicating a greater (and positive) willingness to participate again than those in the long-horizon condition. There is a significant simple effect of horizon when uncertainty was low, with participants in the short-horizon/low-uncertainty cell indicating a significantly higher willingness to participate again than participants in the long-horizon/low uncertainty cell (F = 2.47, p = 0.0594, one-tailed).

The pattern of responses for the questions on task enjoyment and willingness to participate again provide support for the theory that achieving more goals provides more utility. They also confirm a robust finding in the goal setting literature: that working towards and achieving goals enhances interest in the task.

1-5.5 Supplementary analyses

1-5.5.1 Effect of shocks and prior goal achievement on effort

The theory in this study presumes that individual motivation varies with properties of goals that have not yet been achieved. However, it is possible that achievement of (or failure to achieve) *past* goals affected motivation. To test this, I conducted a panel regression with fixed effects of number of tables counted in period T on whether the target was achieved in period T-1, and on the magnitude of the shock at the end of period T-1. A fixed effects model is appropriate because an individual-specific effect, participant ability, is likely correlated with one of the independent variables, whether or not the target was met (Cameron and Trivedi 2009). Participants who were simply faster at counting tables were more likely to achieve a prior period's target. Because participants in the short- and long-horizon conditions faced different numbers of targets, and because the effect of target achievement on motivation might be fleeting, I ran separate regressions for short- and long-horizon participants. In the long horizon, there is only one observation per subject since I am only examining performance following the first target. Thus, I used an OLS regression for participants in the long horizon. However, meeting the target is endogenous and likely to be correlated with ability. Therefore, I used the cumulative shock experienced in periods 1-4 as an instrument for meeting the target. Doing so accounts for participants who just met or just missed the target because of shocks, and captures target achievement due to luck, not ability

The results are shown in Table 1-4 below. Panel A shows the results for participants in the short-horizon conditions. Meeting a prior period's target did not have an effect on effort in the current period. The magnitude of the prior period's shock also had no significant on effort in the current period. Thus, there is no evidence that performance in a prior period had a lasting effect on the subsequent period for participants in the short horizon conditions.

Table 1-4: Effect of shocks and prior goal achievement on effort

Participants counted tables of 1's and 0's in two rounds, a preliminary round and a main round. Each participant earned two points for every correctly counted table. The main round lasted for 40 minutes and was divided into eight 5-minute periods. A random number, or shock, was added to each participant's point total after every 5 minutes. In the *low (high) uncertainty* condition, this random number was uniformly distributed between -2 and 2 (-12 and 12). Participants were also given performance standards that did not affect their pay. *Short (long) horizon* refers to the experimental condition in which participants were given eight 5-minute (two 20-minute) standards of earning 42 (168) points.

Effort refers to the number of points earned from counting tables (not from shocks) during each 5-minute period of the 40-minute main round. Thus, there is a panel of data with 8 observations per subject. *Uncertainty* is 1 for participants in the high-uncertainty conditions. *TargetMet_lag* is an indicator variable that is 1 if the participant achieved a performance standard during the previous 5-minute period, and 0 if the participant failed to do so. *Shock_lag* is an integer indicating the magnitude of the shock received at the end of the previous 5-minute period. *Ability* indicates the number of tables the participant counted during the last 4 minutes of the 10-minute practice round.

Panel A shows a fixed-effects panel regression of *Effort* on *Uncertainty*, *TargetMet_lag*, and *Shock_lag* for participants in the short-horizon conditions. Because each participant had 8 targets, there are 7 opportunities to observe performance following target achievement or failure, and therefore 7 observations per subject. Panel B shows an instrumental variables (2-stage least squares) OLS regression of *Effort* on *Uncertainty*, *TargetMet_lag*, *Shock_lag* ,and *Ability* for participants in the long-horizon conditions, with the cumulative shock experienced through periods 1-4 as an instrument for *TargetMet_lag*. Because each participant had 2 targets, there is only 1 opportunity to observe performance following target achievement or failure, and therefore 1 observation per subject. A '*', '**', or '***' after a coefficient indicate significance at the 10%, 5%, or 1% level, respectively.

Panel A: Effect of shocks and prior goal achievement on effort in short-horizon conditions

	Dependent variable: Effort in period T
TargetMet_lag	0.17
	(0.82)
Shock_lag	0.07
	(0.06)
Constant	42.49***
	(0.49)
Subject fixed effects	Yes
R^2 within	0.0061
R^2 between	0.1528
R^2 overall	0.0203
Num. of obs. (subjects)	469 (67)

Panel B: Effect of shocks and prior goal achievement on effort in long-horizon conditions

	Dependent variable: Effort in period T
Uncertainty	6.26**
	(3.19)
TargetMet_lag	21.67**
	(10.56)
Shock_lag	-0.15
	(0.60)
Ability	0.46
	(0.70)
Constant	20.26***
	(6.07)
R^2 overall	0.3706
Num. of obs.	69

Panel B of Table 1-4 indicates different findings for participants in the long-horizon conditions. Achieving the first target has a very large positive effect on effort in the next five-minute period, even after controlling for ability. Participants earn approximately 14 more points (i.e. counted seven extra tables) in that period as a result of target achievement in the prior period. Magnitude of shock received does not matter. Also, the effect of uncertainty differs from the short-horizon. In the long-horizon, participants facing high uncertainty worked *harder* than those in the low-uncertainty. This may have been driven by the large number of participants in the long-horizon, low-uncertainty condition who appear to have given up and just browsed in the *Leisure Zone*; this explanation will be examined in the next supplementary analysis. Another possibility is that the relative lack of payoffs in the long horizon (i.e. the lower number of available goals) led to boredom. Uncertainty may have alleviated the boredom and induced participants to work harder, as did achieving the first target.

Thus, it appears that achieving or failing to achieve a goal has no effect on future motivation for short-horizon participants, but does have an effect on long-horizon participants. This finding provides weak support for a mechanism operating in parallel with the one proposed in this paper. The next supplementary analysis provides evidence for both explanations.

1-5.5.2 Analysis of decision to enter or forgo the Leisure Zone

Participants faced a choice between work and leisure. Instead of counting tables, they could enter the *Leisure Zone* and browse entertaining content from the internet. Figure 1-11 shows when participants chose to enter the *Leisure Zone*. The 40-minute main round was divided into 5-minute periods; this division is a natural choice because point shocks occurred every 5 minutes. Each 5-minute period was then subdivided into 30-second intervals. Within each 30-second interval, I counted the number of participants who chose to enter the *Leisure Zone*. Figure 1-11 shows this graphically. First, notice how relatively few participants chose to enter the *Leisure Zone*. Second, notice how those that did tend to enter at the beginning of each 5-minute period. This would correspond to the time immediately following a shock, and for participants in the short-horizon conditions, to the time immediately following a goal evaluation.

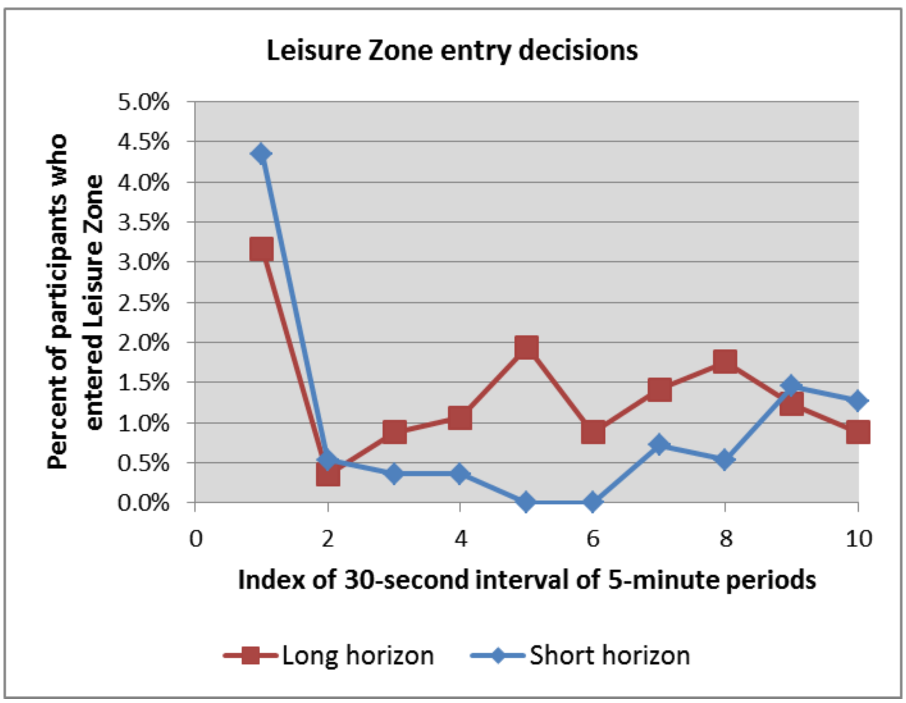


Figure 1-11: Decision to enter the *Leisure Zone* at different points in time. The 40-minute main round was divided into 5-minute periods; this division is a natural choice because point shocks occurred every 5 minutes. Each 5-minute period was then subdivided into 30-second intervals. Within each 30-second interval, I counted the number of participants who chose to enter the *Leisure Zone*. The vertical axis shows the percent of participants in the experimental conditions who chose to enter the *Leisure Zone*.

To investigate this behavior further, I conducted Probit panel regressions of the decision to enter the *Leisure Zone* on participants' observable state, such as the number of points they had accumulated at that second and whether they had recently achieved or failed to achieve a goal, and on a projection of whether they were likely to achieve their current goal if they continued working at the same rate. These regressions are shown below in Table 1-5.

Table 1-5: Factors affecting decision to enter Leisure Zone

Participants counted tables of 1's and 0's in two rounds, a preliminary round and a main round. Each participant earned two points for every correctly counted table. The main round lasted for 40 minutes and was divided into eight 5-minute periods. A random number, or shock, was added to each participant's point total after every 5 minutes. In the *low (high) uncertainty* condition, this random number was uniformly distributed between -2 and 2 (-12 and 12). Participants were also given performance standards that did not affect their pay. *Short (long) horizon* refers to the experimental condition in which participants were given eight 5-minute (two 20-minute) standards of earning 42 (168) points. At any time, participants could stop counting tables and instead enter the *Leisure Zone*, a screen at which they could view web content, such as comic strips and celebrity gossip.

The 40-minute main round was divided into 5-minute periods; this division is a natural choice because point shocks occurred every 5 minutes. Each 5-minute period was then subdivided into 30-second intervals. Within each 30-second interval, I counted the number of participants who chose to enter the *Leisure Zone*. This forms a panel with 80 observations per subject (40 minutes x two 30-second intervals per minute).

Enter Leisure Zone is an indicator variable equal to 1 if the participant chose to enter the Leisure Zone within a 30-second interval. *Time remaining* refers to the number of minutes remaining at the time the decision was made to enter or not enter the Leisure Zone. *Delta* is the difference between the number of points the participant is projected to have at the end of the period, if he continues to count at the current rate, and the target. *PriorGoalMet* is an indicator variable equal to 1 if the participant achieved the goal in the prior period. *ShockLag* is the amount of the point shock experienced at the end of the prior 5-minute period. *CumPoints* is the participants current (cumulative) point total at the time the decision was made to enter or not enter the Leisure Zone.

Panel A shows a Random-effects Probit panel regression of decision to enter the Leisure Zone on experimental conditions, time remaining, and projected target achievement. Panel B shows an OLS regression of Effort on Uncertainty, *TargetMet_lag*, *Shock_lag*, and *Ability* for participants in the long-horizon conditions. Because each participant had 2 targets, there is only 1 opportunity to observe performance following target achievement or failure, and therefore 1 observation per subject.

PANEL A: Random effects Probit panel regression of decision to enter the Leisure Zone on experimental conditions, time remaining, and projected target achievement

	Dependent variable: Entered Leisure Zone
Horizon	0.14
	(0.18)
Uncertainty	0.16
	(0.17)
Horizon x Uncertainty	-0.32
	(0.23)
Time remaining	0.01*
	(0.008)
Delta	-0.004**
	(0.001)
Constant	-2.70***
	(0.13)
Wald χ^2_5	11.53
Prob > χ^2	0.0418
Num. of obs. (subjects)	11,120 (139)

PANEL B: Random effects Probit regression of decision to enter the Leisure Zone on uncertainty and prior goal achievement, short-horizon participants only

	Dependent variable: Entered Leisure Zone
Uncertainty	0.24
	(0.38)
Prior goal met	-0.84**
	(0.34)
Constant	-2.04***
	(0.38)
Wald χ^2_2	6.49
Prob > χ^2	0.0389
Num. of obs. (subjects)	414 (69)

PANEL C: Random effects Probit regression of decision to enter the Leisure Zone on uncertainty and prior shock, and current points total, long-horizon participants only

	Dependent variable: Entered Leisure Zone
Uncertainty	0.38
	(0.24)
Shock_lag	-0.03*
	(0.02)
Cumulative points	-0.004*
	(0.002)
Constant	-1.80***
	(0.22)
Wald χ^2_3	8.83
Prob > χ^2	0.0317
Num. of obs. (subjects)	497 (71)

Panel A of Table 1-5 shows a random effects Probit panel regression of the decision to enter the *Leisure Zone* at each 30-second interval, regressed on the experimental conditions, the time remaining to achieve the current goal, and the *delta*, the projected point shortfall at the end of the current horizon. *Delta* is computed as the number of points the participant is projected to have at the end of the current goal horizon, minus the target. Thus, if the participant is likely to achieve the goal, *Delta* will be positive. This significant and negative coefficient on delta indicates that the less likely a participant is to achieve his goal, the more likely he is to enter the *Leisure Zone*. The positive and significant coefficient on *Time remaining* indicates that the further along participants were in the goal horizon, the less likely they were to enter the *Leisure Zone*. Consistent with these effects being psychological effects that are unaffected by goal horizon and uncertainty, none of the experimental conditions were significant.

Panel B of Table 1-5 shows a random effects Probit panel regression of the decision to enter the *Leisure Zone* at each 30-second interval for participants in the short-horizon conditions only. The significant and negative coefficient on *PriorGoalMet* indicates that participants who achieved the prior period goal were much less likely to enter the *Leisure Zone*.

Panel C of Table 1-5 shows a random effects Probit panel regression of the decision to enter the *Leisure Zone* at each 30-second interval for participants in the long-horizon conditions only. The negative and significant coefficient on lagged shock indicates that participants who received a positive shock were less likely to enter the Leisure Zone; this indicates that the closer they were brought to their target, the less likely they were to quit. The negative and significant coefficient on cumulative points indicates that as participants accumulated more points, they were less likely to enter the *Leisure Zone*.

The analyses in this section indicate that participants' decision to enter the Leisure Zone depends on how likely they are to achieve their goal. As they earn more points and get closer to their targets, they were less likely to enter the Leisure Zone (for many participants, entering the

Leisure Zone is equivalent to giving up on their current target). These findings are consistent with the theory in this paper, which assumes that participants like achieving goals and are willing to exert effort to do so, even if goal achievement is not rewarded with money. However, the finding that participants were less likely to enter the Leisure Zone after achieving a *past* goal suggests a different mechanism, such as affect. Thus, the data from this experiment are consistent with the idea that having a goal motivates, but also consistent with the idea that achieving a past goal elates people and that elation lasts for some time.

1-5.5.3 Other goals

This paper asserts that participant behavior was affected by goals. It is therefore possible that goals other than those assigned by the experimenter affected behavior. Other likely sources of goals are participants' predetermined goals (set before coming to the lab) and goals set by participants during the experiment. Debriefing questions were used to capture these extraneous goals.

To measure participants' predetermined goals, a debriefing question asked "*Did you have a goal of earning a certain amount of money before you came to the lab today?*" Fifty-two percent of participants answered yes. The percentage varied from 50-54% across experimental conditions, but the differences were not significant (in an untabulated ANOVA, neither the main effects of horizon and uncertainty nor their interaction were significant). Of those participants who came to the lab with a goal, 84% said their goal was to earn $16; the others gave responses that ranged from $10-20. Since the predetermined goals did not vary across experimental condition, it is unlikely they affected behavior systematically during the experiment.

To measure goals that participants self-set during the experiment, a debriefing question asked "*Besides the targets that you were given, did you set any goals for yourself during the experiment?*" This question was followed by one that asked participants to describe their goals. Sixty-eight percent of participants reported setting an additional goal during the experiment. The percentages ranged from 61-73% across cells, but an untabulated ANOVA reveals no statistically significant main effects or interaction between horizon and uncertainty. The open-text responses reveal that many participants attempted to increase their table-counting accuracy and a few set higher or lower standards for themselves. However, so few participants reported setting higher or lower standards that the conclusions of this paper are unaffected.

1-6. Concluding remarks

This paper examines how the horizon of performance standards affects worker motivation. I conduct an experiment that shows that performance standards with short horizons can be especially motivating for workers, but that the beneficial effect of short horizons is reduced by uncertainty in the worker's effort-output relationship. The interaction between horizon and uncertainty occurs because workers with short-horizon standards have (and perceive) little control over standard achievement. Despite their effort, negative realizations of uncertainty might prevent them from attaining their standards, and given the short horizon, they do not have time to adjust their effort level in response to random shocks as they occur. Participants who faced short horizons and high uncertainty understood that they had little control over standard achievement. The results are consistent with the theory that workers internalize the performance standards assigned to them in the form of personal goals, and that psychological models of the motivating effect of goals can therefore be applied to performance standards.

My study has several implications. First and foremost, I demonstrate that the horizon of a performance standard can have a measureable impact on worker motivation, even if the level of the standard is held constant. By demonstrating that worker motivation increases when performance standard horizons are properly matched to the uncertainty in production functions, I provide a potential explanation for the variation observed in performance standard horizons in practice.

Next, I show that consistent with past research, workers are likely to internalize performance standards as goals. Practitioners should therefore realize that goals have motivational effects on workers, and that these effects will operate independently of links to extrinsic incentives that are based on standards. Two robust findings in the goal setting literature are that people work harder when working towards a goal, and that goals can make boring tasks fun (Locke and Latham 2002). Therefore, simply providing a standard to workers can motivate them, and managers can augment monetary compensation with satisfaction from achieving goals. This occurs routinely in nonprofit organizations, who are able to attract workers who are willing to accept less compensation in order to pursue their social goals (Hallock 2000).

The motivational benefits of goals are not a free lunch. Since people experience a sensation of loss if they fail to achieve a goal, they will be less motivated if they do not believe they can achieve their assigned goal (Locke and Latham 1990; Heath et al. 1999). To avoid this, managers must choose performance standards carefully. Many previous studies have shown that the level of a performance standard is important. The present results show that the horizon of a standard can have positive effects on motivation, but only when properly matched to the level of uncertainty in the environment. An implication of the interaction between horizon and uncertainty is that investments in accounting systems that reduce uncertainty can give managers more flexibility in choosing the horizon of performance standards.

This study contributes to the managerial accounting literature by demonstrating that the perceived controllability of a performance standard varies with the interaction between the standard's horizon and the uncertainty in the worker's production function, and that controllability affects worker effort. These findings are consistent with the textbook definition of the controllability principle.

My study contributes to the goal setting literature by clarifying earlier mixed findings. Proponents of short-horizon goals had argued that proximity to a goal increases motivation, while opponents argued that short-horizon goals reduce flexibility, lead to failures to achieve goals, and therefore reduce motivation. My study is the first to provide direct evidence of this tradeoff. I demonstrate that uncertainty is the moderating factor that determines whether short-horizon goals are more effective than long-horizon goals.

This study is subject to several limitations. First, the experiment is predicated on the assumption that the performance standards assigned to participants would be internalized and would cause participants to set personal goals. This assumption appears to be valid, as debriefing responses indicate that participants paid attention to their assigned standards. However, some participants set additional goals that need to be analyzed further to determine whether the goal of achieving the external standard (and not another goal) truly drove behavior. Another limitation is that the results may vary when overall performance and achievement of standards have strong effects on extrinsic incentives. Larger financial incentives may swamp the non-monetary psychological effects of internalized goals. More generally, different methods of linking standard achievement to incentives may alter how goals are internalized, and how subjects react to hitting or missing standards. Incentives might also alter the degree of the intensity of a worker's desire to achieve his goal, which Locke and Latham (2002) refer to as goal commitment.

I envision two extensions of this work. This experiment took place in a non-strategic setting and was an out-of-equilibrium test of worker behavior. Future research might examine whether principals understand the motivating effects of performance standards and choose the levels and horizons of performance standards to match the uncertainty in the agent's effort-output relationship. Another interesting direction for future research would be to examine the power of incentives that principals offer on performance standards. In an optimal contracting framework, a controllable measure is one that provides any incremental information about the elements under the agent's control (Antle and Demski 1988). Inclusion of an incrementally informative and controllable measure in a contract will improve the total surplus if the measure is properly weighted (Banker and Datar 1989). In this study, as in other studies of goal setting, the weight on goal attainment was determined by an internal psychological process on the part of the agent. It is possible that agents placed too much weight on goal attainment from an optimal contracting perspective, and it may be that the principal would be better off excluding a performance standard from a contract, even if it is incrementally informative about agent effort. Finally, future research might also investigate whether agents respond differently to performance standards that were chosen strategically by the principal to maximize his profits. In such a setting, fairness concerns such as reciprocity and inequity aversion could easily arise and distort agent behavior from what was observed in the non-strategic setting studied here (Rabin 1993; Bolton and Ockenfels 2000).

Other extensions could explore the use of performance standards to mitigate workers' self-control problems. An emerging view in economics is that the modern organization has evolved to help workers deal with their desires to enjoy utility in the present moment, even though doing so entails a greater cost to future utility (Kaur et al. 2010a)[13]. Under this view, penalties for misbehavior and low output are written into contracts to mitigate such effects; workers supply low effort not because they are effort averse, as is commonly assumed in agency models (Lambert 2001), but because the costs and benefits of effort are not contemporaneous. A worker with a self-control problem will overweight the immediate cost relative to the long-term benefit. By providing an immediate penalty for poor present performance, the firm can offset the effects of the self-control problem by providing a commitment device to the worker[14]. Future studies could alter the timing of the costs and benefits of effort and leisure in an experimental setting, and see if agents and principals choose performance standards to mitigate the effects of self-control problems.

---

[13] The tendency to overweight the present relative to the future is known as present-bias. Present-bias can cause time-inconsistent preferences, such as self-control problems. See Frederick et al. (2002) for a review of time-inconsistency and present-bias, and O'Donoghue and Rabin (1999) for a model of self-control problems.

[14] Kaur et al. (2010b) provide compelling evidence that workers realize they are afflicted by present bias. In a field experiment, they show that workers voluntarily choose contracts with penalties for low output, even when they could have chosen a standard piece rate contract that dominates the penalty contract.

1-7. References

Abeler, J., A. Falk, L. Goette, and D. Huffman. 2011. Reference Points and Effort Provision. *American Economic Review* 101 (2):470-492.

Anderhub, V., S. Gächter, and M. Königstein. 2002. Efficient Contracting and Fair Play in a Simple Principal-Agent Experiment. *Experimental Economics* 5 (1):5-27.

Anderson, S. W., H. C. Dekker, and K. L. Sedatole. 2010. An Empirical Examination of Goals and Performance-to-Goal Following the Introduction of an Incentive Bonus Plan with Participative Goal Setting. *Management Science* 56 (1):90-109.

Antle, R., and J. S. Demski. 1988. The Controllability Principle in Responsibility Accounting. *Accounting Review* 63 (4):700-718.

Ariely, D., E. Kamenica, and D. Prelec. 2008. Man's search for meaning: The case of Legos. *Journal of Economic Behavior & Organization* 67 (3-4):671-677.

Bandura, A. 1989. Self-Regulation of Motivation and Action Through Internal Standards and Goal Systems. In *Goal Concepts in Personality and Social Psychology*, edited by L. A. Pervin: Lawrence Erlbaum Associates, 19-86.

Bandura, A., and D. H. Schunk. 1981. Cultivating Competence, Self-Efficacy, and Intrinsic Interest through Proximal Self-Motivation. *Journal of Personality and Social Psychology* 41 (3):586-598.

Bandura, A., and K. M. Simon. 1977. The role of proximal intentions in self-regulation of refractory behavior. *Cognitive Therapy and Research* 1 (3):177-193.

Bandura, A., and R. Wood. 1989. Effect of Perceived Controllability and Performance Standards on Self-Regulation of Complex Decision-Making. *Journal of Personality and Social Psychology* 56 (5):805-814.

Banker, R. D., and S. M. Datar. 1989. Sensitivity, Precision, and Linear Aggregation of Signals for Performance Evaluation. *Journal of Accounting Research* 27 (1):21-39.

Block, M. K., and J. M. Heineke. 1973. Allocation of Effort under Uncertainty - Case of Risk-Averse Behavior. *Journal of Political Economy* 81 (2):376-385.

Bol, J. C., T. M. Keune, E. M. Matsumura, and J. Y. Shin. 2010. Supervisor Discretion in Target Setting: An Empirical Investigation. *Accounting Review* 85 (6):1861-1886.

Bolton, G. E., and A. Ockenfels. 2000. ERC: A theory of equity, reciprocity, and competition. *American Economic Review* 90 (1):166-193.

Cameron, A. C., and P. K. Trivedi. 2009. *Microeconometrics using stata*. Vol. 5: Stata Press College Station, TX.

De Angelis, D., and Y. Grinstein. 2011. Pay for the Right Performance. *Working paper, Cornell University*.

Fehr, E., S. Gachter, and G. Kirchsteiger. 1997. Reciprocity as a contract enforcement device: Experimental evidence. *Econometrica* 65 (4):833-860.

Frederick, S., G. Loewenstein, and T. O'Donoghue. 2002. Time discounting and time preference: A critical review. *Journal of Economic Literature* 40 (2):351-401.

Frederickson, J. R. 1992. Relative Performance Information - the Effects of Common Uncertainty and Contract Type on Agent Effort. *Accounting Review* 67 (4):647-669.

Hallock, K. F. 2000. Compensation in Nonprofit Organizations. *Research in Personnel and Human Resources Management* 19:243-294.

Harackiewicz, J. M., G. Manderlink, and C. Sansone. 1984. Rewarding Pinball Wizardry - Effects of Evaluation and Cue Value on Intrinsic Interest. *Journal of Personality and Social Psychology* 47 (2):287-300.

Heath, C., R. P. Larrick, and G. Wu. 1999. Goals as reference points. *Cognitive Psychology* 38 (1):79-109.

Heckman, J., C. Heinrich, and J. Smith. 1997. Assessing the performance of performance standards in public bureaucracies. *American Economic Review* 87 (2):389-395.

Hsiaw, A. 2010. Goal Bracketing and Self-Control. *Working paper, Harvard University*.

Indjejikian, R. J., and D. Nanda. 2002. Executive target bonuses and what they imply about performance standards. *Accounting Review* 77 (4):793-819.

Jain, S. 2009. Self-Control and Optimal Goals: A Theoretical Analysis. *Marketing Science* 28 (6):1027-1045.

Kachelmeier, S. J., B. E. Reichert, and M. G. Williamson. 2008. Measuring and motivating quantity, creativity, or both. *Journal of Accounting Research* 46 (2):341-373.

Kahneman, D., and A. Tversky. 1979. Prospect Theory - Analysis of Decision under Risk. *Econometrica* 47 (2):263-291.

Kaur, S., M. Kremer, and S. Mullainathan. 2010a. Self-Control and the Development of Work Arrangements. *American Economic Review* 100 (2):624-628.

Kaur, S., M. Kremer, and S. Mullainathan. 2010b. Self-Control at Work: Evidence from a Field Experiment. *Working paper, Harvard University*.

Kerr, S., and S. Landauer. 2004. Using stretch goals to promote organizational effectiveness and personal growth: General Electric and Goldman Sachs. *Academy of Management Executive* 18 (4):134-138.

Kirschenbaum, D. S. 1985. Proximity and Specificity of Planning - a Position Paper. *Cognitive Therapy and Research* 9 (5):489-506.

Kirschenbaum, D. S., L. L. Humphrey, and S. D. Malett. 1981. Specificity of Planning in Adult Self-Control - an Applied Investigation. *Journal of Personality and Social Psychology* 40 (5):941-950.

Koch, A. K., and J. Nafziger. 2009. Motivational Goal Bracketing. *Working Paper, Institute for the Study of Labor*.

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

Latham, G. P. 2004. The motivational benefits of goal-setting. *Academy of Management Executive* 18 (4):126-129.

Libby, R., R. Bloomfield, and M. W. Nelson. 2002. Experimental research in financial accounting. *Accounting Organizations and Society* 27 (8):777-812.

Locke, E. A. 2001. Self-set Goals and Self-efficacy as Mediators of Incentives and Personality. In *Work motivation in the context of a globalizing economy*, edited by M. Erez, U. Kleinbeck and H. Thierry. Mahwah, N.J.: L. Erlbaum Associates, x, 439 p.

Locke, E. A., and G. P. Latham. 1990. *A theory of goal setting & task performance*. Englewood Cliffs, N.J.: Prentice Hall.

Locke, E. A., and G. P. Latham. 2002. Building a practically useful theory of goal setting and task motivation. A 35-year odyssey. *American Psychologist* 57 (9):705-717.

Locke, E. A., L. M. Saari, K. N. Shaw, and G. P. Latham. 1981. Goal Setting and Task-Performance - 1969-1980. *Psychological Bulletin* 90 (1):125-152.

MacDonald, K. 2011. *The results from estimation commands display only two-sided tests for the coefficients. How can I perform a one-sided test?* StataCorp LP, July 2011 [cited November 23 2011]. Available from http://www.stata.com/support/faqs/stat/oneside.html.

Merchant, K. A. 1985. *Control in business organizations*. Boston: Pitman.

Merchant, K. A., and L. Ferreira. 1988. *HCC Industries*. HBS No. 189-096 ed. Boston: Harvard Business School Publishing.

Murphy, K. J. 2000. Performance standards in incentive contracts. *Journal of Accounting & Economics* 30 (3):245-278.

O'Donoghue, T., and M. Rabin. 1999. Doing it now or later. *American Economic Review* 89 (1):103-124.

Parker, S. C., Y. Belghitar, and T. Barmby. 2005. Wage uncertainty and the labour supply of self-employed workers. *Economic Journal* 115 (502):C190-C207.

Rabin, M. 1993. Incorporating Fairness into Game-Theory and Economics. *American Economic Review* 83 (5):1281-1302.

Read, D., G. Loewenstein, and M. Rabin. 1999. Choice bracketing. *Journal of Risk and Uncertainty* 19 (1-3):171-197.

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

Suvorov, A., and J. van de Ven. 2008. Goal Setting as a Self-Regulation Mechanism. *Working paper, Centre for Economic and Financial Research at New Economic School*.

Tressler, J. H., and C. F. Menezes. 1980. Labor Supply and Wage Rate Uncertainty. *Journal of Economic Theory* 23 (3):425-436.

U.S. Office of Personnel Management. 2011. *Developing Performance Standards* 1998 [cited November 29 2011]. Available from http://www.opm.gov/perform/articles/118.asp.

1-8. Appendix A – Instructions for preliminary round

This appendix contains instructions for the preliminary round. These instructions were displayed one page at a time using a viewer application written in the C# programming language. A snapshot of the viewer application is shown in Figure 1-12 below. The instructions were displayed in a panel on the left, and a sample screen, if available for that page of the instructions, was shown in a separate panel on the right. The panel at right was also used to administer comprehension check questions for the preliminary round.

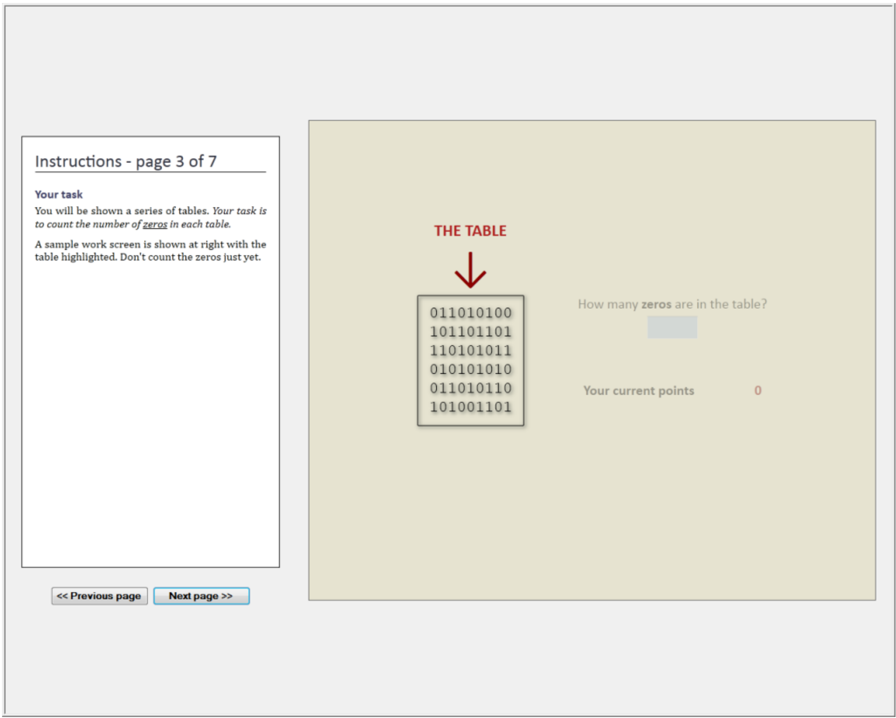


Figure 1-12: Snapshot of viewer application for reading instructions and comprehension check questions. Participants read instructions in the panel on the left, and viewed sample screens, if available for that page of instructions, in the panel on the right.

Instructions - page 1 of 7

Overview

We need your undivided attention for the next 75 minutes. If you have other commitments, please contact the lab administrator so that you can reschedule.

Reading these instructions

Use the [Next page >>] and [<< Previous page] buttons below to read these instructions.

Instructions - page 2 of 7

Restrictions

No electronic devices

You may not use any electronic devices, such as cell phones, PDA's, iPhones, iPads, and laptops. Please turn off or mute your electronic devices at this time. If you cannot agree to this restriction, then you may not participate.

Computer use

You have been assigned a lab computer. You may only use it for the task described below. If you use your assigned computer for any other purpose, you will be asked to leave and will not receive any payment.

Instructions - page 3 of 7

Your task

You will be shown a series of tables. *Your task is to count the number of zeros in each table.*

A sample work screen is shown at right with the table highlighted. Don't count the zeros just yet.

Instructions - page 4 of 7

Points

You will receive 2 points for each table you count correctly.

If you submit *three* incorrect answers for a table, 2 points will be deducted from your point total and a new table will be generated.

Example

You count six tables correctly, and miscount one table three times. You will be awarded points as follows:

Correctly counted tables:	6 x	2 points =	12 points
Penalty:	1 x	–2 points =	–2 points
			10 points

Tracking your points

The work screen will show you the total number of points accumulated. See highlighting at right.

Instructions - page 5 of 7

Your pay

- You will be paid a fixed fee of $10.
- In addition, you will be paid 15¢ for every ten points that you earn.

Instructions - page 6 of 7

How to enter your answers

After you have counted the number of zeros in a table, enter your answer into the blue box and press the *Enter* key on the keyboard.

If you enter the correct answer, ***Correct*** will appear under the blue box and a new table will be generated. If your input was wrong, ***Wrong*** will appear under the blue box, and you will have two additional tries to enter the correct number.

Try it now

Try entering different numbers into the blue box on the sample screen and pressing the *Enter* key. The correct answer is 24.

Click the [Next page >>] button below when you are done with the sample screen.

Instructions - page 7 of 7

If you have questions...

please ask the lab administrator at this time.

When you are comfortable with these instructions...

click the green button below that says *I'm done with the instructions.* You will be asked some questions that test your understanding of these instructions. Once you answer them correctly, your task will begin.

[I'm done with the instructions]

Feel free to use the [<< Previous page] and [Next page >>] buttons to review the instructions while you are answering the questions.

1-9. Appendix B – Instructions for main round

This appendix contains instructions for the main round. These instructions were displayed one page at a time using a viewer application written in the C# programming language. Comprehension check questions for the main round were administered using the viewer application. A snapshot of the viewer application is provided in Figure 1-12 in the previous appendix. Sample screens were shown to participants as part of the main round instructions. The sample screens for pages 4-8 of these instructions are shown in Figure 1-1, Figure 1-4, Figure 1-6, Figure 1-5, and Figure 1-7, respectively.

Different main round instructions were required for each experimental condition. To implement this, I created master document with embedded tags, such as the "**TT**" and "**MM**" tags shown in page 2 of the instructions below. At runtime, the viewer application parsed the master document and replaced each tag with the appropriate value for the experimental condition for that participant. The table below shows the value of each tag for each experimental condition:

Tag	Short-horizon Low uncertainty	Short-horizon High uncertainty	Long-horizon Low uncertainty	Long-horizon High uncertainty
TT	42	42	168	168
MM	5	5	20	20
NT	8	8	2	2
MAX	2	12	2	12
NS	1 random event	1 random event	4 random events	4 random events
OLDRAND	""	""	and any past random events	and any past random events
LONGTEXT	""	""	The window will also let you know how many minutes remain to achieve your target.	The window will also let you know how many minutes remain to achieve your target.

Instructions - page 1 of 9

Now that you have some experience with the task, there are some changes.

What has <u>not</u> changed

- Your task remains the same – counting zeros.
- You will earn 2 points for every table that you count correctly.
- You will lose 2 points if you enter three incorrect answers for a table.
- You will be paid 15¢ for every ten points that you receive.

Click the [Next page >>] button to find out what *has* changed.

Instructions - page 2 of 9

Targets

- People sometimes find it useful to have targets to work towards as they perform tasks like this one. Therefore, to help you perform better ...
- **Your target is to earn {TT} points every {MM} minutes.**
- The rest of the experiment will last for 40 minutes. That means that you will have {NT} targets to achieve.
- Each target is independent.
  - ❑ If you exceed a target, those extra points will *not* count towards the next target.
  - ❑ However, all points earned will count towards your pay.
- There is **no** bonus or penalty for achieving a target or failing to achieve it. *<u>Your pay depends only on the total points that you accumulate.</u>*

Instructions - page 3 of 9

Random Events

- Unexpected events can affect work performance. To depict that here, a random number will be added to – or subtracted from – your point total after every 5 minutes.
- **You will gain or lose up to {MAX} points every time a random event occurs.**
- Since you have {MM} minutes to achieve each target, you will experience {NS} per target.
- The random events can affect target achievement. For example, say that you have exactly {TT} points at the end of a {MM}-minute period. If you then lose 1 point from a random event, you will not achieve your target.

Instructions - page 4 of 9

Revised task screen

The screen has been revised to show progress toward your targets.

- **Target** is your current target.
- **Your current points** shows progress towards your current target. This number includes points earned by counting tables {OLDRAND}.
- **After the next random event** shows how your point total could change after the next random event.
- **Time remaining for this target** is shown in the lower right of the screen. It will be updated every minute.
- Example
  - ❑ Type 24 into the blue box and press *Enter* to see how the numbers change.

Instructions - page 5 of 9

Notification of random events

When a random event occurs, a window will appear to let you know the amount of the gain or loss, as well as your new point total. {LONG-TEXT}

The [Continue] button will appear after the notification window has been open for ten seconds. You should click this button when you are ready to resume work.

Instructions - page 6 of 9

If you fail to achieve a target

When time for a target runs out, the notifications window will let you know if you failed to achieve your target.

After you have finished working on one target, you will start working on the next target. The points and timer shown on your task screen will be reset. Don't worry, all points earned towards previous targets will count towards your pay.

Instructions - page 7 of 9

If you achieve a target

When time for a target runs out, the notifications window will let you know if you achieved your target.

The fireworks will appear every time you achieve a target.

Instructions - page 8 of 9

The *Leisure Zone*

At any time you need a break, you may click the [Leisure Zone] button on the task screen.

When you do, you will be presented with content from a variety of popular websites, such as funny pictures, comic strips, and celebrity gossip. You may view as much content as you wish, but you won't earn points while doing so.

Use the [Next >>] button to see new content. When you are ready to resume counting tables, click the [Show task screen] button.

A sample Leisure Zone screen is shown at right with some sample content displayed. Click the [Next >>] button a few times to see more sample content.

Instructions - page 9 of 9

If you have questions...

please ask the lab administrator at this time.

When you are comfortable with these instructions...

click the green button below that says *I'm done with the instructions*. You will be asked some questions that test your understanding of these instructions. Once you answer them correctly, your task will begin.

[I'm done with the instructions]

Feel free to use the [<< Previous page] and [Next page >>] buttons to review the instructions while you are answering the questions.

CHAPTER 2
THE DARK SIDE OF FEEDBACK

How frequent feedback about goal progress can reduce motivation in the presence of uncertainty

2-1. Abstract

Prior studies have shown that providing performance feedback improves worker performance (Kluger and DeNisi 1996) and imply that more feedback information is better. However, recent evidence suggests possible adverse effects of high-frequency feedback on worker performance (Lurie and Swaminathan 2009; Lam et al. 2011), such as reduced learning and fixation on recent feedback. This paper finds another adverse effect: motivation to achieve a goal decreases when feedback about goal progress is given frequently and there is high uncertainty in the effort-performance relationship. This study uses an experiment with a real-effort task in which participants were rewarded for achieving experimenter-set performance goals. Frequency of feedback about goal progress and uncertainty were each manipulated at two levels, low and high, and the pattern of realizations of uncertainty was also varied. When feedback frequency and task uncertainty were high, participants exerted less effort. The reduction in effort was due to participants' choice to abandon their goal and instead engage in leisure, especially when early realizations of uncertainty were negative and early feedback indicated low performance. These findings imply a benefit to aggregating information: when goal progress information is aggregated over a longer horizon, some of the task uncertainty cancels itself out, and workers are less likely to give up on their goals.

2-2. Introduction

A primary role of managerial accounting systems is to provide information that improves workers' abilities to make organizationally desirable decisions, align the incentives of workers and managers, and motivate workers (Sprinkle 2003). Information from accounting system is often used to set explicit goals, such as budgets, and anticipating receipt of such information, workers may choose to set their own goals (Anand 2012; Sprinkle 2003). Accounting information is also used by workers to assess their progress towards their goals and performance standards (Anand 2012; Heckman et al. 1997). Since goal progress information will affect workers' motivation to achieve their goals (Carver and Scheier 1998), a natural question that arises is how often accounting information about goal progress should be provided to workers.

As an example, consider a worker who operates microchip manufacturing equipment who will receive a large year-end bonus if he improves yield (percentage of chips manufactured that meet certain performance criteria) by 5% from the previous year's level. How frequently should the worker's yield be reported to him in order to motivate him? At one extreme, the worker could be given no feedback during the year. He would only learn whether he achieved his goal at the end of the year. At that point it is too late to make adjustments to the machinery, and he must base all of his choices throughout the year on ex ante information. At the other extreme, the worker might be given daily performance reports on yields. However, if there is any uncontrollable variability in daily yields (e.g. temperature, humidity, and dust inside the facility), daily yields might fluctuate wildly. In that case, the worker would have little idea whether his choice of operating parameters is working without aggregating past yield data. What's worse, if there are a few weeks where the uncontrollable variation is mostly negative, the worker might conclude that his efforts are doomed to fail and abandon his goal. However, if progress is assessed less frequently, the net effect of the uncontrollable variability is likely to be close to its mean (by law of large numbers) and the worker will be better able to infer from the performance report whether his efforts to improve yield are on schedule, and whether the annual goal is still achievable. Anticipating these effects, a manager might choose to provide feedback to the worker less frequently. If she provides feedback to the worker less frequently, the worker will receive less timely information and may take longer to catch problems that have arisen with the machinery. However, if she provides feedback more frequently, and the worker happens to receive early negative feedback, the worker may wish to abandon his yield goal and redirect or reduce effort.

In this study, I use an experiment with a real-effort task to show that motivation to achieve a goal *decreases* when feedback about goal progress is provided frequently and uncertainty in the effort-performance relationship is high. Under these conditions, participants who experienced early negative realizations of uncertainty concluded that their goal was unachievable and abandoned their goal. When uncertainty was lower, the effort-performance relationship was less variable and reported performance more closely reflected effort and participants only abandoned their goal when it was too difficult for them or when they had slacked off early in the goal horizon. When facing infrequent feedback, uncertainty mattered less, because goal progress information was aggregated over a longer period of time and uncertainty realizations cancelled themselves out to a greater degree.

In the experiment, 78 participants performed a computer-based task, counting the number of zeros in a sequence of tables of 1's and 0's. They were awarded a variable number of points for each table counted correctly. In each period, the sequence of point awards differed. Each participant experienced one period in which point awards in the first 5 minutes were mostly low, while point awards in the last 5 minutes were mostly high. In another period, early point awards were mostly high while late point awards were mostly low. And in the other of the three periods, point awards cycled between low, medium, and high values. Participants were assigned a point target for each of three 10-minute periods. Participants were paid extra for each period in which they exceeded their target. The experiment employed a $2 \times 2 \times 3$ mixed design, varying feedback frequency (frequent, infrequent) and task uncertainty (low, high) between-subjects. The sequence of uncertainty realizations (i.e. point awards) were varied at three levels within-subjects, as described above. Participants were provided goal progress feedback at fixed intervals while working. To vary feedback frequency, half of the participants received feedback at shorter intervals than others At each feedback report, participants were given the option of abandoning their task and instead engaging in leisure by browsing the web for pay. However, browsing the web meant giving up on an unachieved target. To vary task uncertainty, the variance of participants' point awards was varied while the mean was held constant.

Results indicate that the number of tables counted in the frequent feedback, high uncertainty condition was lower than that in the other three conditions. This finding is consistent with the hypothesis that motivation decreases when frequent feedback about goal progress is provided to a worker engaged in an uncertain task. This effect was driven by participants' decision to abandon their goal of achieving their target and instead engage in leisure. This is consistent with the hypothesis that early feedback in uncertain environments can lead workers to conclude that their rate of progress is insufficient to achieve their goals and therefore abandon the goals. These two results – that effort is lowest when feedback is frequent and uncertainty is high, and that the reduction in effort is due to goal abandonment – are strongest when participants experienced early negative realizations of uncertainty. Supplemental analysis indicates a possible role of affect, whereby participants whose reported progress is ahead of (behind) schedule increase (decrease) effort following the feedback report.

This paper makes three contributions. First, this paper shows that frequent feedback can be demotivating in highly uncertain environments. This can occur when negative realizations of uncertainty happen to occur in a cluster somewhat early during a worker's goal horizon. The worker then concludes that the goal is unachievable and abandons the goal. A manager, anticipating this behavior, should consider providing goal progress information less frequently, since uncertainty has a tendency to cancel itself out over longer horizons (as per the law of large numbers). Thus, this finding has implications for performance reporting by managerial accounting systems.

Second, this paper contributes to the literature in accounting on the incentive value of information aggregation. Papers in this stream of the literature (see, e.g. Arya et al. 2004) demonstrate that it may be optimal to delay information release to provide optimal incentives to an agent, who has an incentive to shirk if an interim report would indicate either extremely good performance (the agent can then coast) or extremely bad performance (the agent gives up). This paper is also consistent with other papers in this stream that show another benefit to aggregation, that errors in individual items can cancel out (Grunfeld and Griliches 1960; Lim and Sunder 1991; Datar and Gupta 1994).

Finally, the paper contributes to a recent stream of literature in psychology on feedback frequency. This stream has attempted to rebut the view that feedback has universally positive effects and that more feedback is better (Kluger and DeNisi 1996). Recent work has demonstrated an adverse effect of frequent feedback on learning and task performance, because of the cognitive cost of processing feedback (Lam et al. 2011), and that frequent feedback coupled with high uncertainty can reduce decision-making performance because of fixation on recent feedback (Lurie and Swaminathan 2009). This paper adds to that literature by demonstrating that frequent feedback coupled with high uncertainty can reduce motivation to achieve a goal.

The remainder of the paper is organized as follows. Section 2-3 discusses prior literature and develops hypotheses. Section 2-4 discusses my research decision. Section 2-5 discusses my results, and section 2-6 concludes.

2-3. Theory and hypotheses

2-3.1 Feedback

There is a large body of literature in accounting and in psychology on feedback and its effects on task performance. A common assumption that has emerged from this literature is that feedback improves performance, and in many cases is a necessary condition for performance (Kluger and DeNisi 1996; Locke and Latham 2002). For example, in their review of the findings in the goal setting literature, Locke and Latham (2002) state that goals are only effective when feedback about goal progress is provided. However, in their review of the effects of feedback on performance, Kluger and DeNisi (1996) show that the results of many feedback studies are not as one-sided as is commonly assumed. They describe numerous examples of studies whose findings indicate that feedback does not always have positive effects. They describe moderating factors, such as the sign of feedback, the frequency of feedback, and presence of a goal, that affect the relationship between feedback and performance. Additionally, when given feedback, workers must stop to process it. Such processing is cognitively taxing (Lam et al. 2011), and the interruption can reduce productivity. Thus, recent evidence suggests possible adverse consequences to providing feedback to workers. This study adds to that literature by demonstrating another possible adverse effect of frequent feedback – worker motivation to achieve a goal can decrease when feedback about goal progress is provided frequently and task uncertainty is high.

2-3.2 Processing of and response to feedback on goal progress

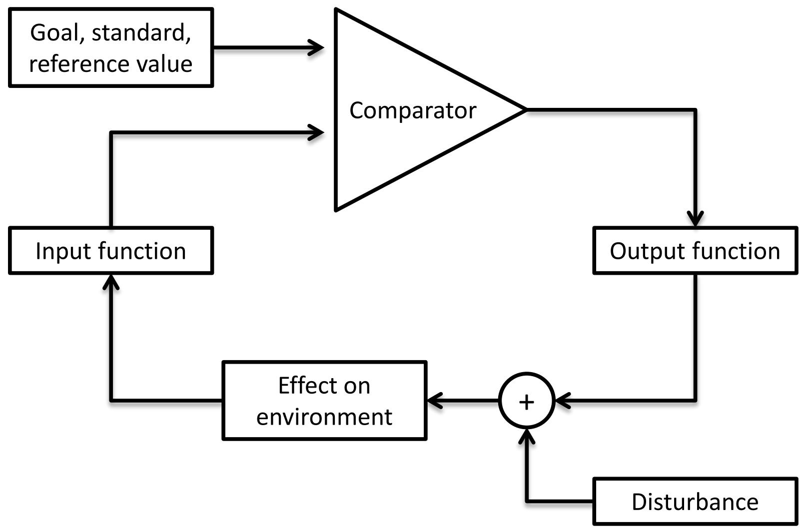


Figure 2-1: Schematic depiction of a feedback loop, the basic unit of cybernetic control. In such a loop a sensed value is compared to a reference value or standard, and adjustments are made in an output function (if necessary) to shift the sensed value in the direction of the standard. Source: Carver and Scheier (1998), figure 2.1, page 11.

How do individuals process information about their progress towards a goal? Carver and Scheier (1998) model the psychological process that governs goal-directed behavior as an engineering control system (see Figure 2-1). The current state of the system is compared to a reference value, which is a desired state. In a negative-feedback control system, a discrepancy between the current state and the desired state changes behavior to reduce the discrepancy. In this study, the behavior of interest is the rate of effort exertion.

When an individual is given feedback about *progress* toward a goal, the individual does not compare current progress to the final goal. That would be like a salesman comparing his first quarter sales to her annual sales goal and concluding that the goal has not been achieved[15]. Instead, individuals compare current performance with performance expected at that point in time (e.g. comparing Q1 performance to one fourth of the annual goal, or to an a priori expectation of Q1

[15] Prior literature has examined the satisfaction that individuals derive from failing to achieve a goal, or from performance that exceeds a goal. See Heath et al. (1999) for a thorough review that models goal utility using the Prospect theory value function (Kahneman and Tversky 1979). Instead, this paper focuses on responses to information on *goal progress* that is provided while time remains to achieve the goal.

performance). In effect, individuals compare the *rate* of progress at the time of the feedback report with the rate of progress expected at that point in time (which may or may not be the average rate required to achieve the goal). If the actual rate of progress differs from the desired rate of progress, individuals adjust effort so that the actual and desired rates will match in the future. Expanding the feedback control system analogy, Carver and Scheier (1998) argue that actual progress is analogous to position in a physical system, while the rate of progress is analogous to velocity (the first derivative of position). Continuing the above example, actual Q1 performance is analogous to distance, while the rate of performance (performance per unit time) is analogous to velocity. Carver and Scheier (1998) model the process governing the rate of effort exertion as a velocity meta-controller setting the reference value for the position controller. The meta-controller is checking on how well the position controller is doing at reducing the discrepancies it is trying to reduce. In other words, is progress towards the goal occurring at a desired rate? If not, the rate of effort is adjusted up or down to bring the discrepancy reduction rate to a desired level.

2-3.2.1 Goal achievement and goal abandonment

Carver and Scheier (1998) present evidence that a discrepancy between the actual rate of performance and the desired rate triggers a behavioral response. Individuals who discover that their rate of progress is lower (higher) than desired experience negative (positive) feelings, and increase (decrease) their efforts in response. Thus, negative feelings cause an increase in the rate of effort, and positive feelings cause a decrease in the rate of effort.

When informed of a discrepancy between actual and desired rates of performance, two extreme cases need to be considered. If an individual learns that he has achieved his goal (and assuming there is no benefit to exceeding the goal), he will cease exertion of effort – the rate of effort will drop to zero. The other case is that the individual will conclude that he is so far behind schedule that he cannot achieve the goal within the goal's horizon. Therefore, his best response is to give up on the goal.

2-3.3 A theory of optimal feedback frequency about goal progress

Assume an individual has a goal and must achieve the goal within some time frame (the horizon). The individual can only achieve the goal by exerting effort on some task, and that the individual can adjust his rate of effort over the goal horizon. Task performance is increasing in effort, but is also affected by environmental uncertainty (e.g. $Y = e + \theta$, where Y is task performance, e is effort, and θ is noise). The uncertainty is assumed to have zero mean. If, during the goal's time frame, task performance exceeds some threshold, the goal will be achieved and the individual will experience some reward; the reward may simply be the intrinsic utility from goal achievement, but could also be extrinsic[16]. Finally, assume that the individual cannot determine progress towards the goal without a progress report from some external reporting system. This last assumption is realistic when information must be collected and aggregated through a time-consuming process. For example, consider a manager responsible for warranty costs. Such costs are measured over time and reported periodically. Another example is an executive whose bonus depends on quarterly profit and must wait until all adjusting entries for the quarter are made to learn firm profit. Finally, someone facing a relative performance evaluation scheme must wait until her peers' performance is tallied and reported to discover her own performance.

How will the feedback provided in the progress reports affect the individual's motivation to achieve the goal? How does the magnitude of the noise affect the individual's effort choice? How frequently should the progress reports be provided to the individual? Would a manager, acting strategically, have a preference over reporting frequency? The remainder of this section attempts to answer these questions by providing a theory of the effects of varying the frequency of goal progress information in the presence of uncertainty.

[16] In their review of the goal-setting theory literature, Locke and Latham (2002) note that monetary incentives can be used to enhance goal commitment. In effect, monetary incentives increase the valence of the goal and therefore individuals are willing to exert more effort to achieve the goal. They also note that when goals are very difficult, paying people to achieve the goal can decrease effort because people abandon the goal.

2-3.3.1 Effect of providing discrete feedback on goal performance

Carver and Scheier (1998) distinguish between continuous and discrete feedback. For some tasks, individuals can easily monitor their own performance and have perfect knowledge of their results. For example, a car salesman can easily track the number of cars he has sold in the previous month. In other cases, individuals do not know their performance unless they receive a report from an external system. For example, a plant manager has to compile a report of machine downtime in order to compute the percentage downtime for the previous month. In many accounting settings, feedback is provided to workers periodically, not continuously. How do individuals respond to and process discrete feedback?

Carver and Scheier (1998) note that when feedback is provided discretely, individuals must stop to process the feedback. Such processing is cognitively taxing and can actually reduce performance and learning. Lam et al. (2011) find an inverted-U relationship between performance and feedback frequency. Performance is highest at some intermediate level of feedback frequency, but drops if feedback is provided too frequently because individuals' attention is directed away from the task. Performance also drops if frequency is too low since individuals are not provided with enough information to adjust their input to their task. These works suggest a potential main effect of feedback frequency on effort[17], depending on the chosen levels of feedback frequency.

Another behavior that results when individuals stop to process feedback about progress toward a goal is reassessment of expectancy. In expectancy-value theory (Bandura 1989; Vroom 1964), expectancy is defined as an individual's subjective assessment of the probability that effort will lead to performance. When individuals are provided with new information about their progress towards a goal, they reconsider their goal expectancy – the likelihood that further effort will lead to goal achievement. Assuming a fixed utility from achieving a goal, an individual will reassess the

[17] Another reason why effort might be lower as feedback frequency is increased is self-control. Prior studies (e.g. Kaur et al. 2010) document that present-biased individuals are less likely to exhibit self-control problems when they get into a rhythm with their work. More frequent interrupts from feedback are likely to disrupt workers' rhythms.

optimal level of effort to provide during the remainder of the goal horizon. If the individual is slightly behind schedule, the optimal decision might be to slightly increase effort. The increase in effort will be costly, but may still be worth it, given the expected goal utility. If the individual is far behind schedule, the individual's best response might be to give up on the goal; the additional effort required might be too costly given the expected goal utility. Figure 2-2 depicts this relationship.

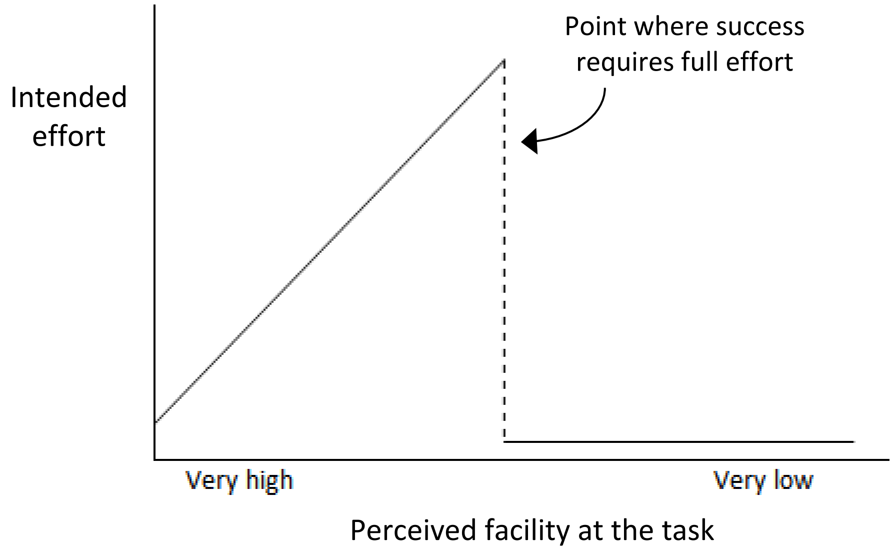


Figure 2-2: Optimal effort as a function of likelihood of achieving the goal. When the likelihood of achieving a goal is very high, very little effort might be required to guarantee goal achievement. As the likelihood of goal achievement decreases, an individual wishing to maximize expected utility might trade off costly effort for higher expected goal utility. However, at some point the effort cost will exceed the expected goal utility and the individual's optimal response is to give up on the goal and exert zero effort. Adapted from Kukla (1972).

There is an alternative behavioral response to feedback indicating that an individual is either behind or ahead of schedule in goal progress. Carver and Scheier (1998) have demonstrated that goal-related feedback indicating more (less) progress than expected creates positive (negative) affect. They then argue that positive (negative) affect leads to a reduction (increase) in effort. A contrasting finding is that of Erez and Isen (2002), who find that positive affect increases motivation and task performance. They specifically test the components of expectancy theory and find that all three, expectancy (subjective belief that effort will lead to performance), instrumentality (subjective belief that performance will lead to a reward), and valence (subjective value of the reward) increase when subjects experience positive affect. Given the contrasting findings on the role of affect on effort, I offer the following two-sided hypothesis:

> **Hypothesis 1**: Effort will change when an individual receives a feedback report indicating unexpected progress toward a goal.

2-3.3.2 Effect of task uncertainty and its interaction with feedback frequency

Uncertainty is likely to affect individuals' assessments of goal expectancy. As the level of uncertainty is increased, realizations of uncertainty are likely to be more extreme. This can be particularly problematic when individuals do not have knowledge of their results and must rely on an external reporting system. When reports of goal progress are provided *infrequently*, uncertainty becomes less of a concern. Many realizations of uncertainty will occur between progress reports, and by the law of large numbers (Weisstein 2013), the average of those realizations should approach the true mean, which is assumed to be zero. In other words, as the interval between feedback reports increases, the sum of a series of zero-mean shocks will go to zero. However, if feedback about goal progress is provided frequently, fewer realizations of uncertainty will have occurred between reports, and the sum of those realizations is more likely to deviate from the true mean. An individual relying on a progress report that contains only a few extreme realizations of uncertainty may draw different inferences than someone who experiences a greater number of aggregated realizations.

Returning to the microchip fabrication worker example from section 2-2, imagine weekly progress reports on an annual yield goal. If there is enough uncontrollable variation in short-term machine performance, the weekly reports will highlight that variation and could easily obscure the true trend in cost (see Figure 2-3). According to Carver and Scheier's theory, the worker will reassess goal expectancy at each progress report. If the worker experiences a few bad weeks in a row, he might conclude that his efforts to improve yield are failing and give up on the goal (see Figure 2-4), or renegotiate it with his superiors. Note that if uncertainty is very low, performance will closely track effort and feedback will be more informative about the effects of effort on performance. If uncertainty is low, the worker will only quit if he learns he cannot exert effort at a rate sufficient for goal achievement.

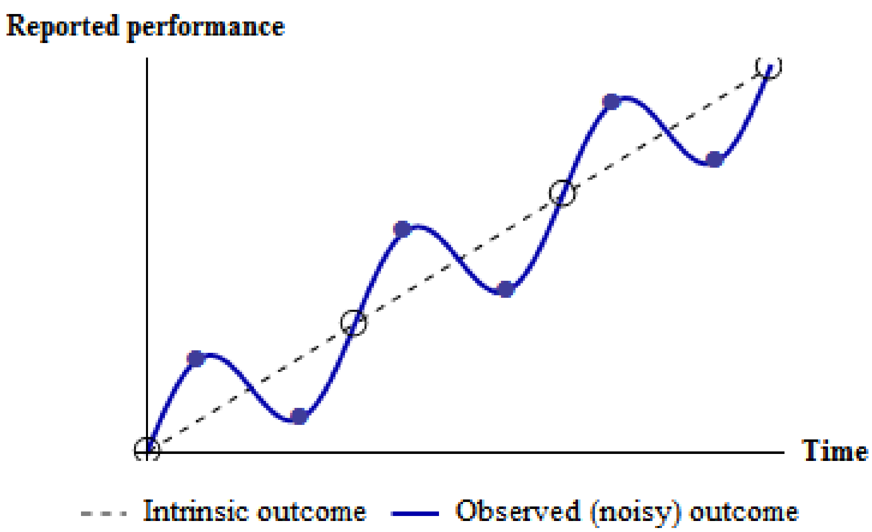


Figure 2-3: Performance as a function of time in a situation where performance is the sum of worker effort plus cyclical noise with zero mean. The worker's rate of effort exertion is assumed to be constant over time, and performance is assumed to be a linear function of effort. If there were no noise, performance as a function of time would be represented by the dashed line. However, because of the noise, observed performance will follow the curved solid line. The dots on the curve represent sampling of performance at different frequencies. The solid dots show performance sampled at twice the frequency of the noise, and the hollow dots show performance sampled at the exact frequency of the noise.

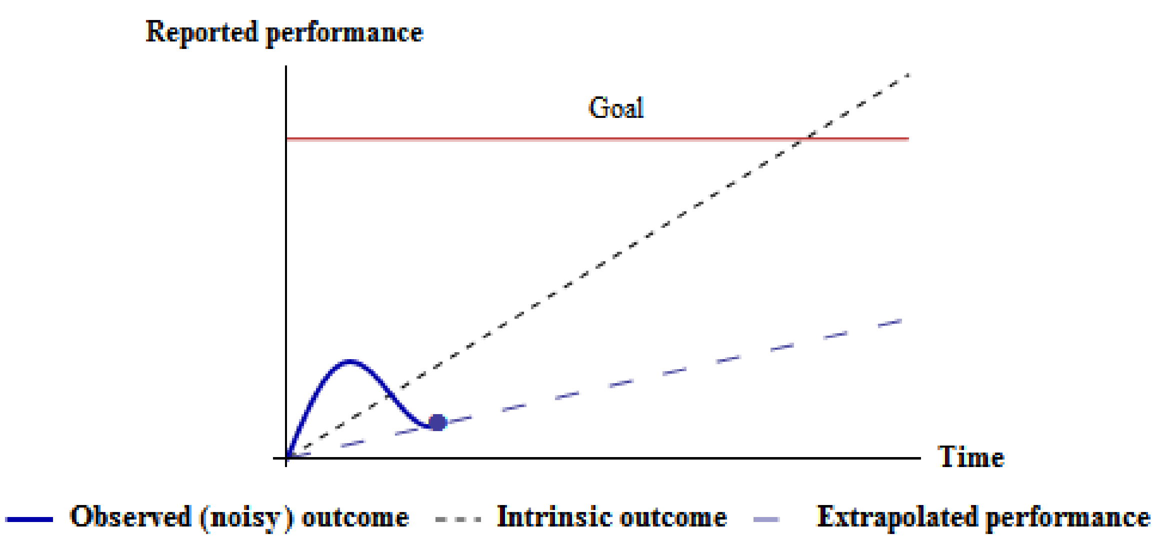


Figure 2-4: Inferences drawn from frequently sampled performance. Assume worker performance equals effort plus zero-mean, cyclical noise, and therefore follows the solid curved line. Further assume the worker does not know the relationship between effort and

performance. If the first sampling of performance occurs at the solid dot on the graph, the worker will infer that the effort-performance relationship follows the lower dashed line, even though the true relationship is the higher dashed line. The worker will extrapolate future performance based on the lower dashed line and conclude that it is not possible to achieve the goal in the allotted time.

2-3.3.3 Value of aggregation of information

The previous section establishes that when frequent feedback about goal progress is provided to a worker facing a noisy task, the worker might conclude, based on early information, that the goal is unachievable and quit. When feedback is provided less frequently, realizations of uncertainty will aggregate, and by the law of large numbers, the sum of these uncertainty realizations should approach their mean, zero, as feedback frequency is reduced[18]. This effect is depicted in Figure 2-5 below.

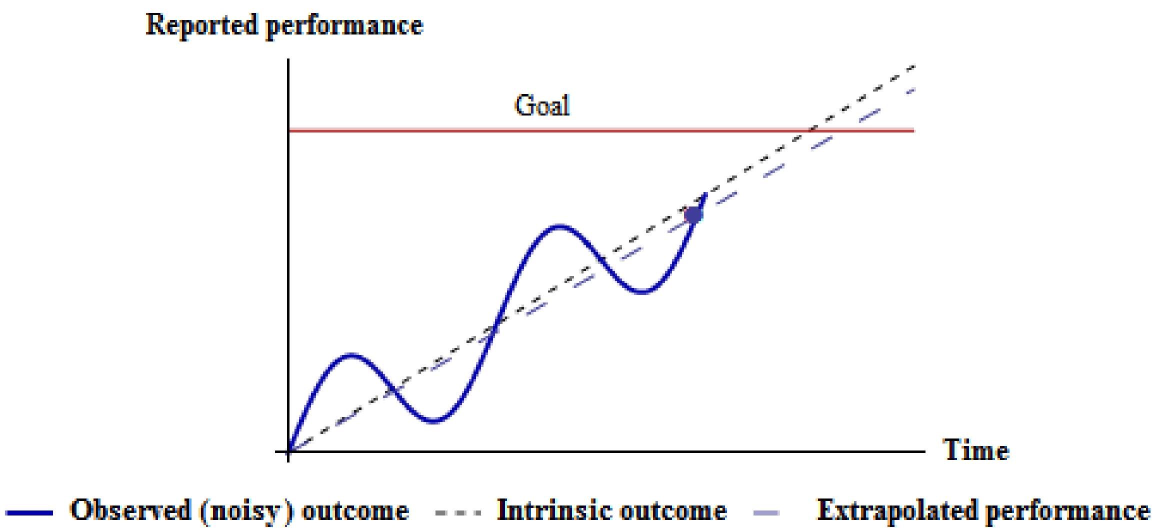


Figure 2-5: Inferences drawn from infrequently sampled performance. Assume worker performance equals effort plus zero-mean, cyclical noise, and therefore follows the solid curved line. Further assume the worker does not know the true relationship between effort and performance (top dashed line). If the first sampling of performance occurs at the solid dot on the graph, the worker will infer that the effort-performance relationship follows the lower

[18] Aggregation of recent performance information in this setting is analogous to a low-pass filter in signal processing. A low-pass filter is a circuit that allows low-frequency signals to pass through and filters (attenuates) high-frequency signals (Siebert 1986). A famous theorem in signal-processing theory states that the minimum sampling rate for alias-free sampling (Aliasing is an effect that causes different signals to become indistinguishable. For example, in Figure 2-3, if sampling occurred only at the hollow dots, it would appear that there is no noise.) is two times the highest frequency component of the signal (Siebert 1986). This rate is known as the Nyquist rate. In this paper I argue that when attempting to measure the performance of a worker attempting to achieve a goal while working on a noisy task, it may be desirable to sample at rate *lower* than the Nyquist rate so as to obscure the effect of the noise.

dashed line, even though the true relationship is the higher dashed line. The worker will extrapolate performance during the allotted time using the lower dashed line. If the worker faces a goal, she is likely to conclude that it is possible to achieve the goal in the allotted time.

A manager who must choose feedback frequency could choose to act strategically and reduce feedback frequency. The manager faces a tradeoff – less frequent feedback is obviously less timely and gives the worker less time to respond to information about progress. On the other hand, the worker might quit earlier than the manager would like if early realizations of uncertainty are negative. The strategic choice of degree of information aggregation inherent in this setting is similar to that modeled by Arya et al. (2004), who model the intertemporal aggregation of information and its incentive value. They model a similar tradeoff as the one I describe here and find that delaying information release can be valuable in providing incentives because workers cannot respond to extreme outcomes that may occur in short time spans. They use a two-period model and show that a worker who is doing either very well or very poorly after one period has less incentive to exert effort in the second period. Hence, the principal may wish to only provide performance information after both periods.

Another possible scenario is that the microchip fabrication worker has a few weeks in a row in which the normal unobservable variation in performance leads to higher yields. The worker might conclude that his efforts are working better than expected and reduce effort on the initiatives to improve yield. Later, when the net effect of uncertainty is better understood, the worker might need to increase effort. However, it may be too late at that point and the worker might abandon the goal. Thus, when uncertainty is high and feedback is provided frequently, clusters of extreme realizations of uncertainty are more likely to cause workers to reduce effort, and to possibly abandon their goals.

The reasoning in this section leads to the following pair of hypotheses:

Hypothesis 2a: Effort will be lowest when feedback is provided frequently and uncertainty is high.

Hypothesis 2b: When feedback is provided frequently and uncertainty is high, effort will be lowest because individuals abandon their goals.

Hypotheses 2a and 2b are consistent with rational behavior and assumptions common to standard economic models, such as maximization of expected utility defined over final wealth states. However, the behavioral assumption that people over-generalize from small samples (see, e.g. Tversky and Kahneman 1971; Rabin 2002) can strengthen the hypotheses. In environments where feedback is provided frequently, people who receive closely-spaced feedback reports may be more likely to overweight recent feedback. Lurie and Swaminathan (2009) find exactly this. They manipulate feedback frequency and uncertainty in a decision-making task[19], and find that individuals excessively focus on recent data and do not appropriately compare information across multiple time periods. In this study, individuals who over-rely on small samples are even more likely to adjust effort and/or abandon their goals after early frequent feedback reports.

2-3.3.4 Sequences of realizations of uncertainty

In any sequence of draws of a random variable, clusters of values in certain ranges are likely to occur. For example, when flipping a coin, it is common to see clusters of heads in any sequence. Based on the previous hypotheses, these sequences are likely to have different effects when feedback frequency and uncertainty are varied. When feedback is provided infrequently, the cumulative effect of these clusters is likely to be reduced. For example, if heads is assigned a value of 1 and tails is assigned -1, the sum of a sequence of flips is likely to be closer to zero in a longer sequence than in a smaller sequence, by the law of large numbers. However, when feedback about the sum of a sequence of flips is provided more frequently, the reported value is more likely to deviate from zero. Thus, in a setting where feedback is provided frequently and realizations of

[19] Participants in Lurie and Swaminathan (2009) work on the news vendor problem, in which they have to choose the optimal number of newspapers to stock. There is a cost to having unsold inventory, but an opportunity cost to too little inventory. Lurie and Swaminathan manipulate the degree of noise and feedback frequency and example profit outcomes. While their manipulations are similar to mine, they examine performance on a decision-making task whereas I test motivation. Additionally, their work does not examine the effect of feedback frequency and uncertainty on performance in the presence of a goal.

uncertainty can either be bad, good, or neutral (e.g. zero effect), goal-related motivation is likely to differ depending on the cluster of realizations experienced. When early realizations are negative, a worker might conclude a goal is unachievable and give up. However, when early realizations are positive, a worker might conclude he is so far ahead of schedule that he can reduce effort. Later, when the sum of uncertainty goes to zero, the worker might find that he did not exert sufficient effort and can no longer achieve the goal. These two possibilities are much less likely when there are no early clusters of uncertainty realizations. I therefore make the following hypothesis:

> **Hypothesis 3**: When feedback is provided frequently and uncertainty is high, effort will be highest and goal abandonment will be lowest when early feedback is consistent with expectations; effort will be lowest and goal abandonment will be highest when feedback is below expectations; both will be moderate when feedback when feedback is above expectations.

2-4. Method

2-4.1 Experimental design and task

2-4.1.1 Overview of experimental task and design

This study used a laboratory experiment to examine the impact of feedback frequency and uncertainty on effort in the presence of a goal. Participants performed a straightforward, repetitive task (same task as in Anand 2012; Abeler et al. 2011) and received monetary compensation if their task performance exceeded a target. The experiment used a 2 × 2 × 3 mixed design. Feedback frequency and uncertainty were manipulated at two levels between subjects. When working on the task, participants were provided feedback about goal progress either provided frequently or infrequently. Task uncertainty (variance in the reward per unit of the task) was either low or high. Each participant participated in three 10-minute periods. Feedback frequency and uncertainty were held constant for each participant across all three periods, but the sequence of realizations of uncertainty differed in each period. For each unit of the task completed, participants were awarded either *Low*, *Medium*, or *High* points (the values assigned to these labels differed based on the uncertainty condition and are described in section 2-4.1.4.2). In one of the three periods, realizations of uncertainty at the beginning of the period were mostly *Low*, and reverted to *High* at the end of the period; this sequence of uncertainty realizations is referred to as the "bad early" sequence since early outcomes hindered achievement of the goal. In another of the three periods, participants experienced the "good early" sequence, which was the opposite of the "bad early" sequence; in the "good early" sequence, early outcomes were mostly *High* and therefore aided in goal achievement. Participants also experienced a period in which the realizations of uncertainty cycled between the three possible values. This was referred to as the "neutral" sequence since at any time during the period, the cumulative effect of uncertainty was always close to zero. The order of the three sequences – good early, bad early, and neutral – was randomized across subjects to prevent order effects.

2-4.1.2 Experimental task

The theory underlying the experiment predicts that motivation will vary with feedback frequency and uncertainty, and that changes in motivation will manifest themselves as changes in effort exerted. Therefore, to measure motivation, I used a real-effort task[20].

[20] In a real-effort task, participants must exert cognitive or physical effort in order to accomplish the task, and can choose how much effort to exert. For examples, see Kachelmeier et al. (2008), who required participants to design rebus puzzles, a cognitively demanding task. Also see Ariely et al. (2008), who had participants assemble structures out of Legos, a task that required physical and cognitive effort. The alternative to a real effort task is one where the experimenters ask participants to choose a number that represents effort. For example, Anderhub et al. (2002) study reciprocity in a principal-agent setting; participants assigned to the role of agents chose a number between 0 and 20 to represent their effort choice. Effort in their experiment was costly in that it reduced agents' payout, but the different levels of effort did not require different levels of cognitive or physical exertion. The practice of choosing a number is quite common in experimental economics (see, e.g. Fehr et al. 1997; Frederickson 1992), but not appropriate in an experimental setting where human motivation is expected to vary.

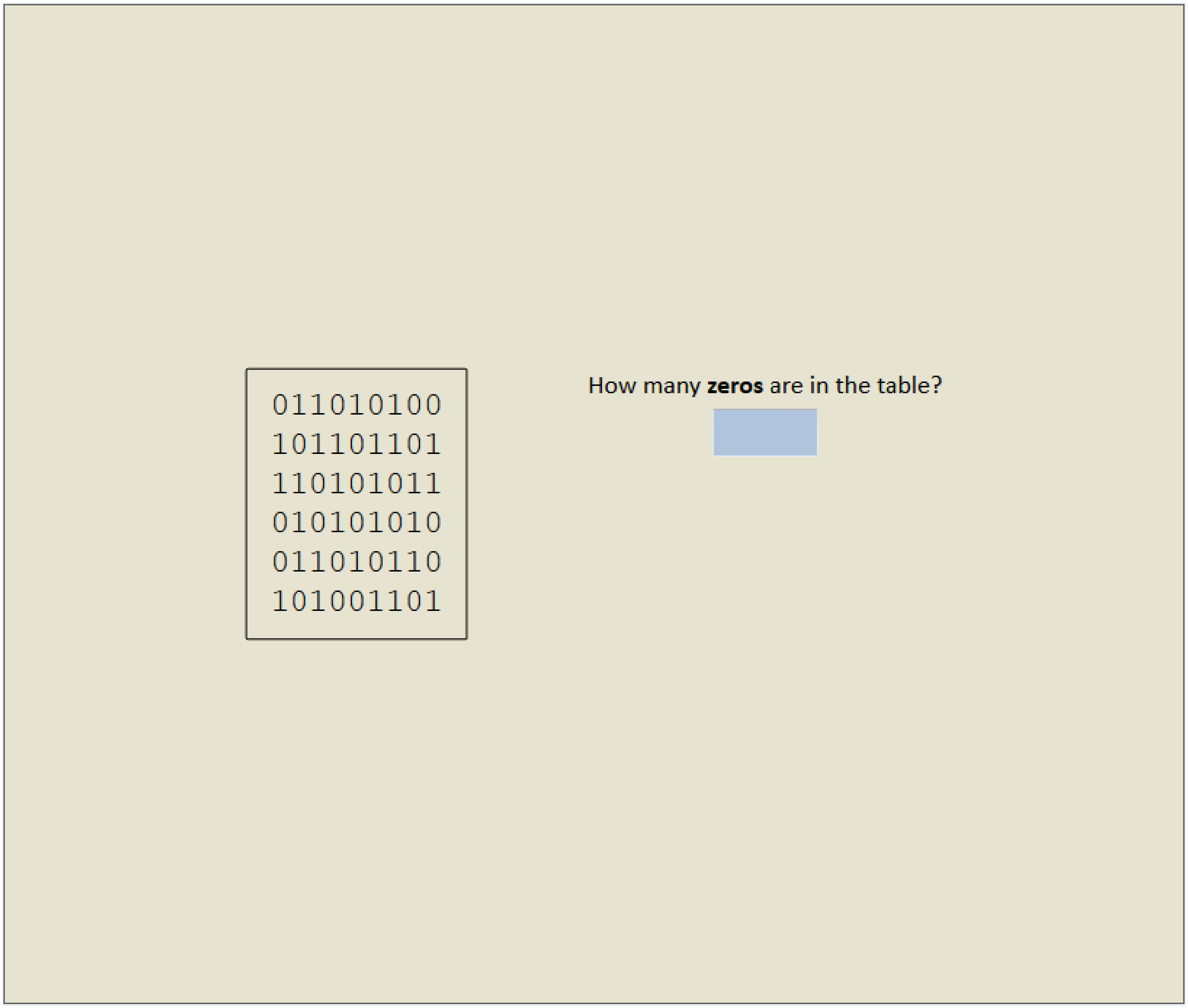


Figure 2-6: The screen on which participants performed their task during the main round of the experiment. The table is the box with 1's and 0's. Participants had to count the number of zeros in the table and enter it into the textbox labeled "How many zeros are in the table?" If they entered a correct answer, the word *Correct* appeared under the textbox, the table flashed, and a new table was displayed. If they entered an incorrect answer, the word "*Wrong"* appeared under the textbox, followed by "Tries left: 2" or "Tries left: 1". Three tries were allowed for each table, and a new table was displayed after three incorrect answers. Participants were awarded points for each table counted correctly, and penalized if three incorrect responses were given for a table. The point total was not displayed during the main round. During the practice round, however, participants' current point total was displayed under the blue box and updated every time a participant finished working on a table.

Participants' effortful task was to count the number of zeros in tables comprised of 54 randomly generated zeros and ones (see Figure 2-6). The task is adapted from Abeler et al. (2011), who argued that it was useful for assessing real effort because it "*does not require any prior knowledge, performance is easily measurable, and there is little learning possibility; at the same time, the task is boring and pointless, and we can thus be confident that the task entailed a positive cost of* *effort for subjects.*" The positive cost of effort is desirable for this study since the theory employs an assumption common to most agency models – that effort is costly (Lambert 2001).

Participants were awarded points for successfully counting tables, and details of the point awards are provided in section 2-4.1.4. Participants were also penalized for providing incorrect responses. Participants were given three tries to count each table, and penalized 10 points after providing three incorrect answers[21]. This scheme is adapted from Abeler et al. (2011), and is designed to discourage guessing without penalizing participants for accidental counting mistakes.

Participants were also given the option to forgo the table-counting task and instead engage in leisure by browsing the web. Every feedback report contained a button with the caption *Leisure Zone*. If a participant chose to click that button, the Microsoft Internet Explorer browser, version 10, was opened and the task screen was hidden. Participants could do anything they wished in the browser. However, once in the Leisure Zone, participants had to remain there until the end of the current 10-minute period. Thus, entering the *Leisure Zone* meant giving up on an unachieved goal. I recorded the details of participants' decision to enter the *Leisure Zone*, including the time at which they chose to enter it as well as their point total at the time they entered.

2-4.1.3 Dependent measures

The dependent variables are effort and abandonment rates. Effort is measured by the number of tables counted correctly in each 10-minute period. Thus, effort is a repeated measure. Abandonment rate is measured by the proportion of participants who abandoned their target and chose to enter the *Leisure Zone*. Participants' decisions to enter the Leisure Zone after earning enough points to meet the target are *not* included in the abandonment measure.

---

[21] As described in section 2-4.1.4, the mean point award for correctly counting a table is 10 points in all experimental conditions.

2-4.1.4 Description of experimental conditions

Participants were assigned a target number of points for each of the three 10-minute periods. The target was the same for each period. The target was specific to each participant and was determined in real-time based on each participant's performance in a 10-minute practice round (to be described in section 2-4.3.1). Participants' pay increased by $5 for each period in which they achieved the target[22]. Each period was independent of the others and participants began each period with zero points.

2-4.1.4.1 Feedback frequency manipulation

During the main round of the experiment, participants counted the number of 0's in tables of 1's and 0's. For each table counted correctly, participants were awarded some points; the point awards appeared random to participants. The computer did *not* inform participants of the point award for each table. Instead, the computer provided feedback about aggregate performance at fixed intervals. Participants in the *frequent feedback* condition received performance feedback every minute, while those in the *infrequent feedback* condition received performance feedback every four minutes. The feedback report informed participants of the number of points they had accumulated since the previous feedback report, as well as the number of points remaining to hit the target. The feedback report also notified participants of the rate at which they had accumulated points since the previous feedback report, as well as the average rate required during the remainder of the period to hit the target. A sample feedback screen from the *frequent feedback* condition is shown in Figure 2-7. If the participant's point total was greater than or equal to the target at the time a feedback report was displayed, the feedback screen notified the participant and

[22] Numerous studies in the goal-setting literature pay participants for goal achievement (Locke and Latham 2002). A conclusion of this literature is that paying for goal achievement enhances goal commitment. However, paying for a very difficult goal can hurt performance is people believe they cannot achieve their goal. This effect is consistent with what I predict theoretically.

displayed animated fireworks. A sample progress report indicating target achievement is shown in Figure 2-8.

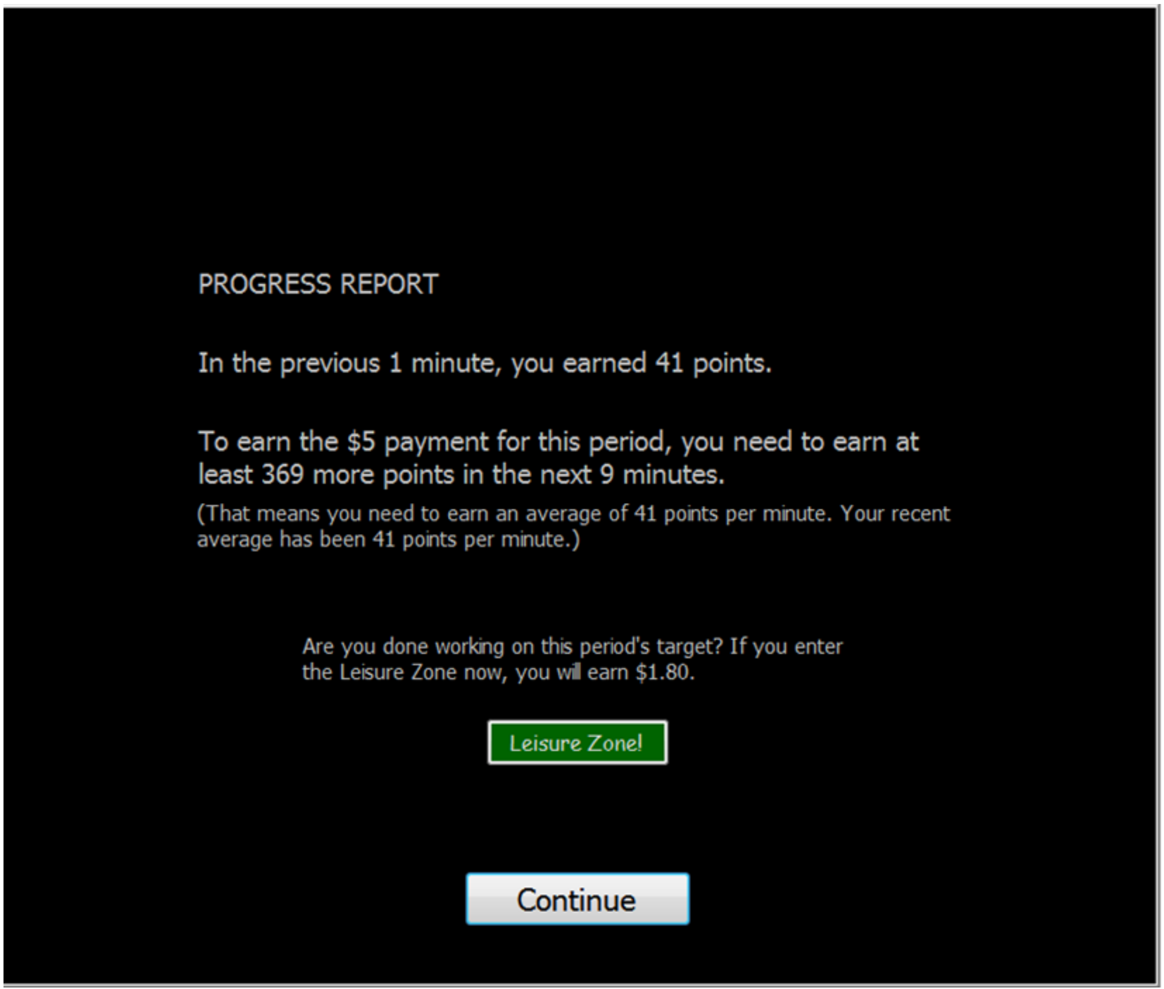


Figure 2-7: A sample feedback screen from the *frequent feedback* condition. This screen was displayed every 1 minute for participants in the *frequent feedback* conditions, and every 4 minutes for participants in the *infrequent feedback* conditions. The screen shows that the participant earned 41 points since previous feedback report or since the beginning of the period, and informs the participant that (s)he must earn at least 369 more points to achieve the target (which is $41 + 369 = 410$ in this case). The text in parentheses compares the rate at which the participant earned points since the last feedback report with the average rate required during the remainder of the period in order to achieve the target. At each feedback report, the participant is also given the option of giving up on the target and entering the *Leisure Zone*, wherein (s)he can browse the web, and earn $0.20 / minute for doing so. Participants can either click the *Leisure Zone* button to browse the web, or the Continue button to resume counting tables. Entering the *Leisure Zone* was an irreversible decision. Upon entering it, participants were not allowed to count tables until the next period.

In all conditions, a final feedback report was provided at the end of each 10-minute period. The feedback report informed the participant whether or not they achieved their target and earned

$5 for that 10-minute period. A sample final feedback screen indicating failure to achieve the target is shown in Figure 2-9.

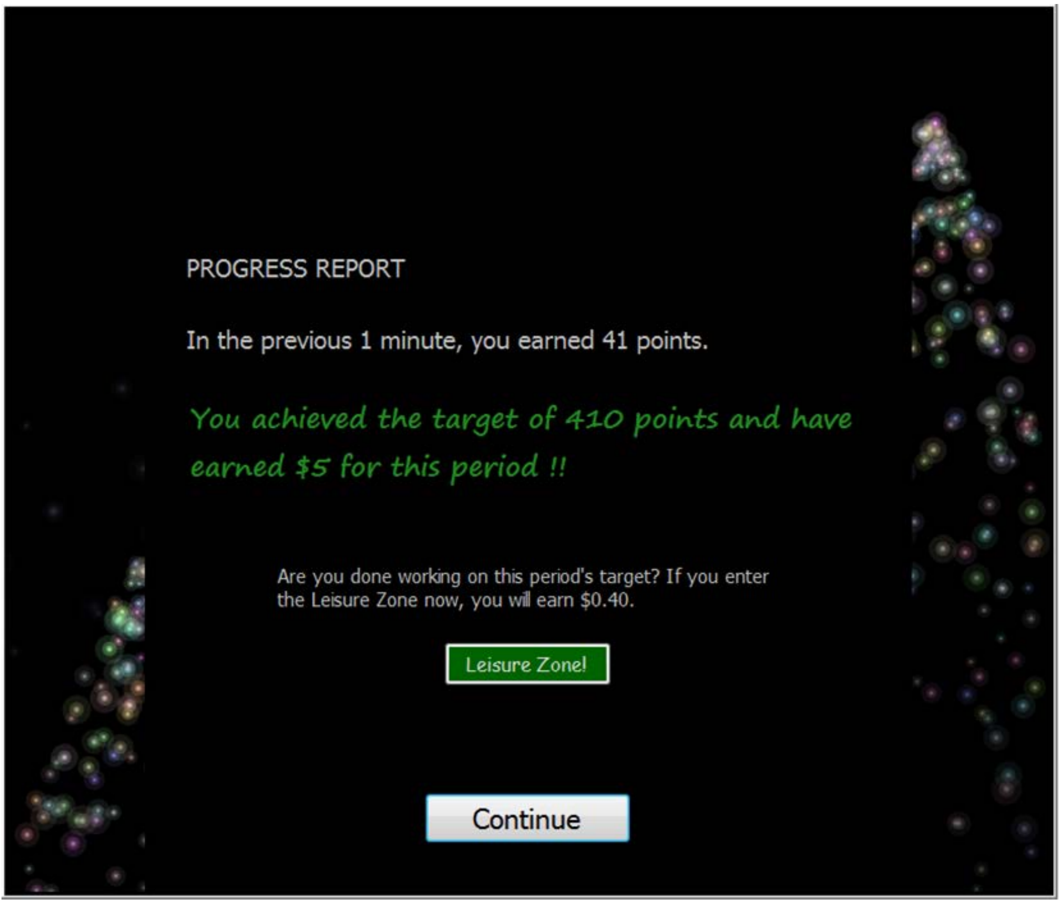


Figure 2-8: A sample feedback screen from the *frequent feedback* condition. This screen was displayed every 1 minute for participants in the *frequent feedback* conditions, and every 4 minutes for participants in the *infrequent feedback* conditions. The screen shows that the participant earned 41 points since previous feedback report and also notifies the participant that (s)he has accumulated at least the target number of points and has earned $5 for the period. This notification was accompanied by animated fireworks. The participant is still given the option of giving up on the target and entering the *Leisure Zone*, wherein (s)he can browse the web, and earn $0.20 / minute for doing so. Participants can either click the *Leisure Zone* button to browse the web, or the Continue button to resume counting tables. Entering the *Leisure Zone* was an irreversible decision. Upon entering it, participants were not allowed to count tables until the next period.

A timeline for each 10-minute period is shown in Figure 2-10.

Each time a feedback screen was displayed, participants had the option of entering the *Leisure Zone*. They could do so by clicking the *Leisure Zone* button displayed on the screen (see Figure 2-7 and Figure 2-8). If a participant chose to do so, a web browser was displayed and the task screen was hidden. The task screen remained hidden until the subsequent 10-minute period.

Thus, entering the *Leisure Zone* implied giving up on the current target, or that the participant had already achieved the target. Participants were paid $0.20 per minute spent in the *Leisure Zone*. The

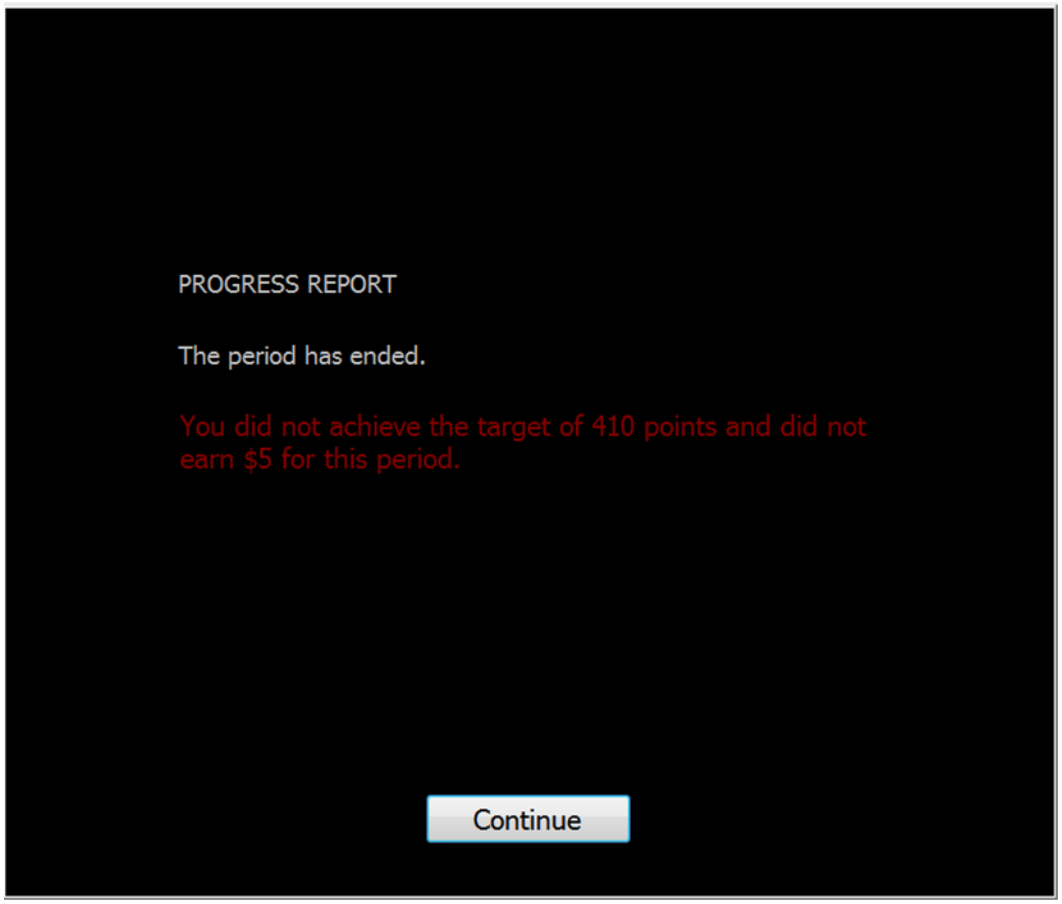


Figure 2-9: A final feedback screen presented at the end of each of the three 10-minute periods. The screen informed participants that the period had ended. It also stated whether they had achieved their target and earned the $5 payment for doing so. Notification of failure was presented in dark red text, and notification of success was presented in dark green text, along with animated fireworks (see previous figure).

feedback screen informed participants how much they would earn if they entered the *Leisure Zone* at that time, and participants could trade off the expected payoff from the goal against money that could be earned in the *Leisure Zone*. Alternatively, they could enter the *Leisure Zone* after learning that they had already achieved the current period's target.

Every time a feedback screen was displayed, participants were interrupted. Since participants in the *infrequent feedback* conditions were only shown two feedback screens during each 10-minute period (at minutes 4 and 8), while those in the *frequent feedback* conditions were shown nine feedback screens during each period (at each of minutes 1 through 9), participants in the *infrequent feedback* conditions were interrupted fewer times. However, since interruptions

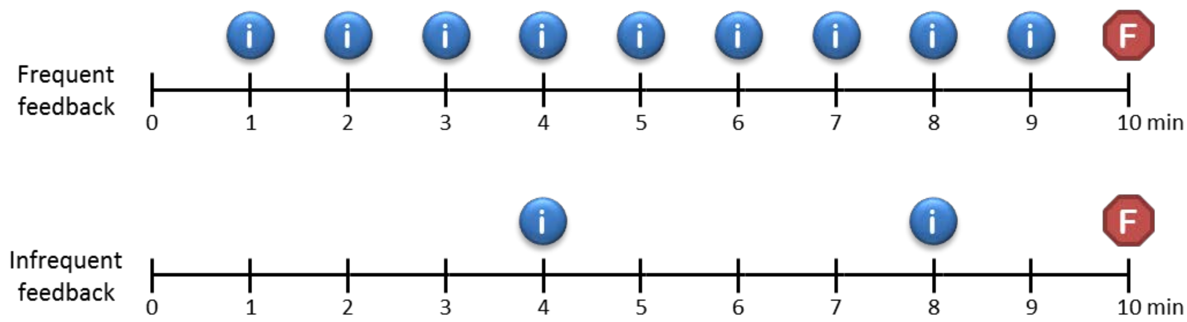


Figure 2-10: Timeline of the task. Participants counted tables of 1's and 0's for three 10-minute periods. This figure depicts the timing of one such period. Participants were given a target number of points for each 10-minute period (the target was the same in each period) and began each period with zero points. Points were awarded for each table counted correctly, and points were subtracted for three incorrect responses on a given table. Participants did not learn the point award for each individual table. Instead, participants only learned their progress at fixed intervals. Sample progress reports are shown in Figure 2-7 and Figure 2-8. In the figure, each symbol represents a feedback report; the feedback reports were displayed every 1 minute in the *frequent feedback* condition, and every four minutes in the *infrequent feedback* condition. At the end of each 10-minute period, a final report notified the participant whether (s)he had achieved the target (see Figure 2-9 for sample report). This is indicated by the symbol.

could potentially occur in the middle of counting a table, participants in the frequent feedback conditions could potentially be at an economic disadvantage since they could lose more of their work. To reduce the economic disadvantage while still preserving the psychological effect of interruptions, I displayed the feedback screen during a 15-second window around the desired event time after a table had been completed. So for example, if a feedback screen was supposed to be displayed at minute 4, I would open a time window at 3 minutes and 52.5 seconds (3' 52.5"). If the participant completed a table by either submitting a correct response or a third incorrect response between 3' 52.5" and 4' 7.5", I would display the feedback screen at that time. If no table completion event occurred between 3' 52.5" and 4' 7.5", the feedback screen was displayed at 4' 7.5". By displaying feedback screens in this manner, the difference in amount of work lost across conditions was hopefully minimized.

2-4.1.4.2 Uncertainty manipulation

Participants were awarded points for each table counted correctly. Participants in the *low uncertainty* condition received 8, 10, or 12 points per table, while those in the *high uncertainty* condition received 2, 10, or 18 points per table. Thus, the mean point award (10 points) was the same in all experimental conditions, but the variance differed[23]. This implementation of uncertainty is intended to represent task uncertainty: each unit of the task completed could have a different outcome per unit of effort.

2-4.1.4.3 Sequences of uncertainty realizations

In any sequence of draws of a discrete random variable, clusters of the same value are likely to occur. The theory underlying this experiment suggests that clusters of either Low or High values can affect behavior. Therefore, to control for their occurrence across subjects, I pre-programmed sequences of draws of point awards. Participants participated in three 10-minute periods. In one of those periods, they experienced mostly *Low* point awards (i.e. 2 or 8 points in the high or low uncertainty conditions, respectively) early in the period and mostly *High* point awards (i.e. 12 or 18 points in the low or high uncertainty conditions, respectively) late in the period; this is referred to as the "bad early" sequence since outcomes that hindered target achievement occurred early. In another period, participants experienced mostly *High* point awards early in the period and mostly *Low* point awards late in the period. This "good early" sequence was the opposite of the "bad early" sequence. The good early and bad early sequences of point awards, which are specified for each minute of each 10-minute period, are detailed in Table 2-1.

In another period, point awards for each table cycled through the *Low*, *Medium*, and *High* values. This was referred to as the "neutral" sequence. In this sequence, point awards cycled through the following list: *High*, *Medium*, *Low*, *Low*, *Medium*, *High*. Thus, at any time during the

[23] The uncertainty manipulation was a mean-preserving spread. This implementation of uncertainty is consistent with prior related experiments (e.g. Anand 2012; Lam et al. 2011), as well as with prior analytical work investigating the effect of uncertainty on labor supply (e.g. Block and Heineke 1973).

period, the cumulative effect of uncertainty was near zero under the neutral sequence; participants' score under this sequence was always close to the mean point award (10 points) times the number of tables completed.

Table 2-1: Pre-programmed sequences of uncertainty realizations

Participants counted the number of zeros in tables of 1's and 0's during three 10-minute periods of the main round. Points were awarded for each table counted correctly. Each point award was *Low*, *Medium*, or *High*. In the *low uncertainty* condition, these values corresponded to 8, 10, and 12 points, respectively. In the *high uncertainty* condition, these values corresponded to 2, 10, and 18 points, respectively. The sequence of point awards was pre-programmed. In one of the 10-minute periods, participants experienced the "bad early" sequence of uncertainty, wherein bad outcomes (i.e. mostly *Low* point awards) occurred early in the period. The sequence reverted to *High* point awards towards the end of the period. The "bad early" sequence is given below in Table 2-1a, and the "good early" sequence is given in Table 2-1b. The tables show the point awards for the first 4 tables counted in each of the 10 minutes, as well as the point awards for any additional tables counted during a minute. The average counting rate was approximately 4 tables per minute. Both sequences perfectly revert: minute 10 is the opposite of minute 1, minute 9 is the opposite of minute 2, etc. Also notice that the "good early" and "bad early" sequences are opposites of each other.

Table 2-1a: Point awards in the "bad early" sequence of uncertainty realizations

Minute	First table counted	Second table counted	Third table counted	Fourth table counted	Subsequent tables counted
1	Low	Low	Low	Medium	Medium
2	Low	Low	Low	Low	Medium
3	Medium	Medium	Medium	Medium	Medium
4	Low	Low	Medium	Medium	Medium
5	Low	Low	Medium	Medium	Medium
6	High	High	Medium	Medium	Medium
7	High	High	Medium	Medium	Medium
8	Medium	Medium	Medium	Medium	Medium
9	High	High	High	High	Medium
10	High	High	High	Medium	Medium

2-4.2 Participants

Seventy-eight students at a large, private research university participated in this laboratory experiment. Participants were recruited through the web site of the business school's laboratory, which maintains a large pool of subjects for use in experimental research. Since the task required no specific skills or knowledge beyond basic computer skills, the ability to exert effort on a straightforward task, and the ability to respond to a typical pay-for-performance arrangement, this subject pool was an appropriate choice (Libby et al. 2002). Upon seeing advertisements for studies

Table 2-1b: Point awards in the "good early" sequence of uncertainty realizations

Minute	First table counted	Second table counted	Third table counted	Fourth table counted	Subsequent tables counted
1	High	High	High	Medium	Medium
2	High	High	High	High	Medium
3	Medium	Medium	Medium	Medium	Medium
4	High	High	Medium	Medium	Medium
5	High	High	Medium	Medium	Medium
6	Low	Low	Medium	Medium	Medium
7	Low	Low	Medium	Medium	Medium
8	Medium	Medium	Medium	Medium	Medium
9	Low	Low	Low	Low	Medium
10	Low	Low	Low	Medium	Medium

at the lab's web site, members of the subject pool voluntarily sign up for experiments. This study's advertisement stated "*In this experiment you will be asked to perform a computer-based task for money. We expect that the study will take 60 minutes on average.*" I deliberately withheld details of the compensation scheme, such as the expected average payment, from the recruiting materials so as to reduce the likelihood that participants came to the lab with monetary goals.

2-4.3 Procedures

Upon entering the lab, each participant was seated at a computer station with partitions on both sides. The participant was assigned to a treatment condition using a predetermined random-assignment scheme, and verbally instructed to "*read and sign the consent form, if you agree, and then follow the instructions on the screen*." The experimental session was divided into four periods: a

10-minute practice round and a main round consisting of three 10-minute periods. The computer provided instructions before the practice round and before the each period of the main round. The computer also administered comprehension and manipulation checks after each set of instructions, as well as debriefing questions after the main round.

2-4.3.1 Practice round

After signing the consent form, participants viewed interactive instructions for the practice round (see appendix in section 2-8). These instructions described the rules governing the task and incentives. After reading the instructions, the computer presented participants with comprehension check questions. These tested participants' knowledge of the task, points awarded per table, rate at which points will be exchanged for cash, and the show-up feed for the experiment. The program required correct answers on all questions before the participant could proceed. On all questions, more than 88% of participants answered the questions correctly on the first try. On only one instance did someone provide two incorrect responses on the same question. Thus, it appears that participants understood the instructions.

During the practice round, participants counted the number of 0's in tables of 1's and 0's and entered their count into a textbox (see screen snapshot in Figure 2-6). Participants received exactly 10 points for each table counted correctly. Participants were penalized 10 points if they provided three incorrect responses on a given table, and then presented with the next table. *The computer screen reported participants' current point total while they were working in the practice round* (see Figure 2-6), but participants were not aware that the round's duration was 10 minutes, and no timer was displayed. After the practice round ended, the software informed participants of the number of tables they had counted correctly and their pay for the practice round. Participants earned 8¢ per point accumulated during the practice round and were paid at the end of the experiment. After the practice round, participants were shown a screen that informed them of the amount of money they had earned so far. This amount was the $5 show-up fee, plus $0.08 per point earned in the practice round. After dismissing this screen, participants progressed to the main round of the experiment. The practice round gave participants an opportunity to learn the task before the main round. Data collected during the practice round also provides a baseline measure of variation in individual performance.

2-4.3.2 Target setting

Anand (2012) found significant individual variation in table counting speeds using the exact same task. That implies that, for a given target, some individuals will find it so easy to achieve that it provides little incentive to work. Other people will find the same target impossible to achieve in the time allotted, and that target will provide no incentive for those people. To remedy this, I assigned an individual-specific target. (Anand 2012) found that table counting speeds significantly improved during his 10-minute practice round, and that counting speeds improved a little more during his main round. Using his data, I computed individual-specific targets as follows. I used the number of tables counted by a participant in the second half of the 10-minute practice round, and doubled it. This gives an estimate of the number of tables the participant would count if he worked at that rate for a full 10 minutes. I then multiplied by 10 points per table, the average point award per table. I then increased that by 17.5%, which was approximately the improvement that Anand (2012) found from his practice round to his main round. I rounded that number to the nearest 10. I used a lower bound of 410 points, and an upper bound of 550 points. Thus, if c represents the number of tables counted in the second half of the practice round, the target would be:

$$Target = Max\left(Min\left(\left[Round\left(\frac{2c \times 10\,\frac{\text{points}}{\text{target}} \times 1.175}{10}\right) \times 10\right], 550\right), 410\right)$$

2-4.3.3 Main round

After completing the practice round, participants were shown instructions for the main round, which are reproduced in the appendix in section 2-9. These instructions informed participants that they would continue to perform the same task of counting tables of 1's and 0's.

However, participants were informed that they would work for three 10-minute periods, and that they would be given a target for each period. They were further informed that they could increase their pay by $5 for each target achieved, up to a maximum of $15 if targets were achieved in all three periods. Participants were told that failure to achieve a target in a given period carried no penalty, but also meant that no payment would be received for points accumulated during that period. Participants were informed that each period was independent of the others and that they would begin each period with zero points.

The instructions then described the uncertainty and feedback frequency manipulations (described above), as well as the Leisure Zone. Participants were made aware that they would not be provided a running total of their points, and would only learn their progress when the feedback window appeared. Participants were shown samples of the screens that they would see (Figure 2-6 through Figure 2-9) while reading the instructions. They were also notified that once they entered the Leisure Zone, they could not resume counting tables until the next period. Participants were also notified that they could earn $0.20 per minute spent in the Leisure Zone.

After reading the instructions, the computer presented participants with comprehension and manipulation check questions. These tested participants' knowledge of the instructions. The program required correct answers on all questions before the participant could proceed. On all questions, more than 85% of participants answered the questions correctly on the first try. Only 78% of participants correctly answered the question about the frequency at which the feedback window would appear on the first try. However, on only four instances did someone provide two incorrect responses on the same question. Thus, it appears that participants understood the instructions.

2-4.3.3.1 Three periods of the main round

Participants worked for three 10-minute periods. In each period, they experienced a different sequence of realizations of uncertainty, the good early, bad early, and neutral sequences.

The order of sequences was randomized across participants. At the end of each period, participants were shown a screen indicating whether they had achieved their target or not and earned the $5 for that period (Figure 2-9). Before each period, participants were shown a screen that said "*It is time to begin the (first / second / third) period. You will begin this period with zero points. If you earn [target] or more points this period, your pay will increase by $5. When you are ready to start period (1 / 2 / 3), click the Continue button below*."

2-4.3.3.2 Debriefing and final payment

After completing the 30-minute main round, participants were shown a screen that informed them of the total pay they had earned from the $5 show-up fee, achieving targets, and from time spent in the *Leisure Zone*. Participants then answered 15 debriefing questions about their motivation, feeling of control over target achievement, emotion experienced from the progress reports, reasons for giving up on a target, interruptions, self-set goals, and task enjoyment. After completing these questions, the software instructed the participant to go to the lab's office, where the experimenter paid the participant in cash[24]. The participant then left the lab.

2-4.4 Software

The task, instructions, comprehension check questions, and debriefing questions were administered using custom software written in the C# programming language. The software was designed to allow participants to perform the experiment with minimal involvement from the experimenter. The software captured every keystroke and mouse movement made by each participant, and used Microsoft's Internet Explorer, version 10, for the *Leisure Zone*.

[24] After each participant completed debriefing questions, the software sent the experimenter an email notifying him of the participant's computer number and final pay. This allowed the experimenter to remain in the office in case many participants finished at the same time and had to line up to be paid.

2-5. Results

2-5.1 Cell means of dependent variables

The two dependent variables are effort and goal abandonment. Effort is measured by the number of tables counted correctly in each 10-minute period of the main round. Goal abandonment is measured by the proportion of participants who abandoned their target and chose to enter the *Leisure Zone* instead. Figure 2-11 and Figure 2-12 depict the mean of the effort and giving up dependent measures, respectively, in each experimental condition. Table 2-2 reports cell means for these two dependent measures. As predicted, effort, as measured by the number of tables counted, is lowest in the high-frequency, high-uncertainty experimental condition. This is true regardless of sequence of uncertainty realizations. Figure 2-11 and Table 2-2 also indicate that effort is lowest for the bad early sequence of uncertainty, except in the frequent feedback, low uncertainty experimental condition. Also, Figure 2-12 and Table 2-2 show that, as predicted, the proportion of participants who abandoned is generally highest in the frequent feedback, high uncertainty condition; the only exception is for the good early sequence of uncertainty, where the rate of giving up is highest in the frequent feedback, low uncertainty condition.

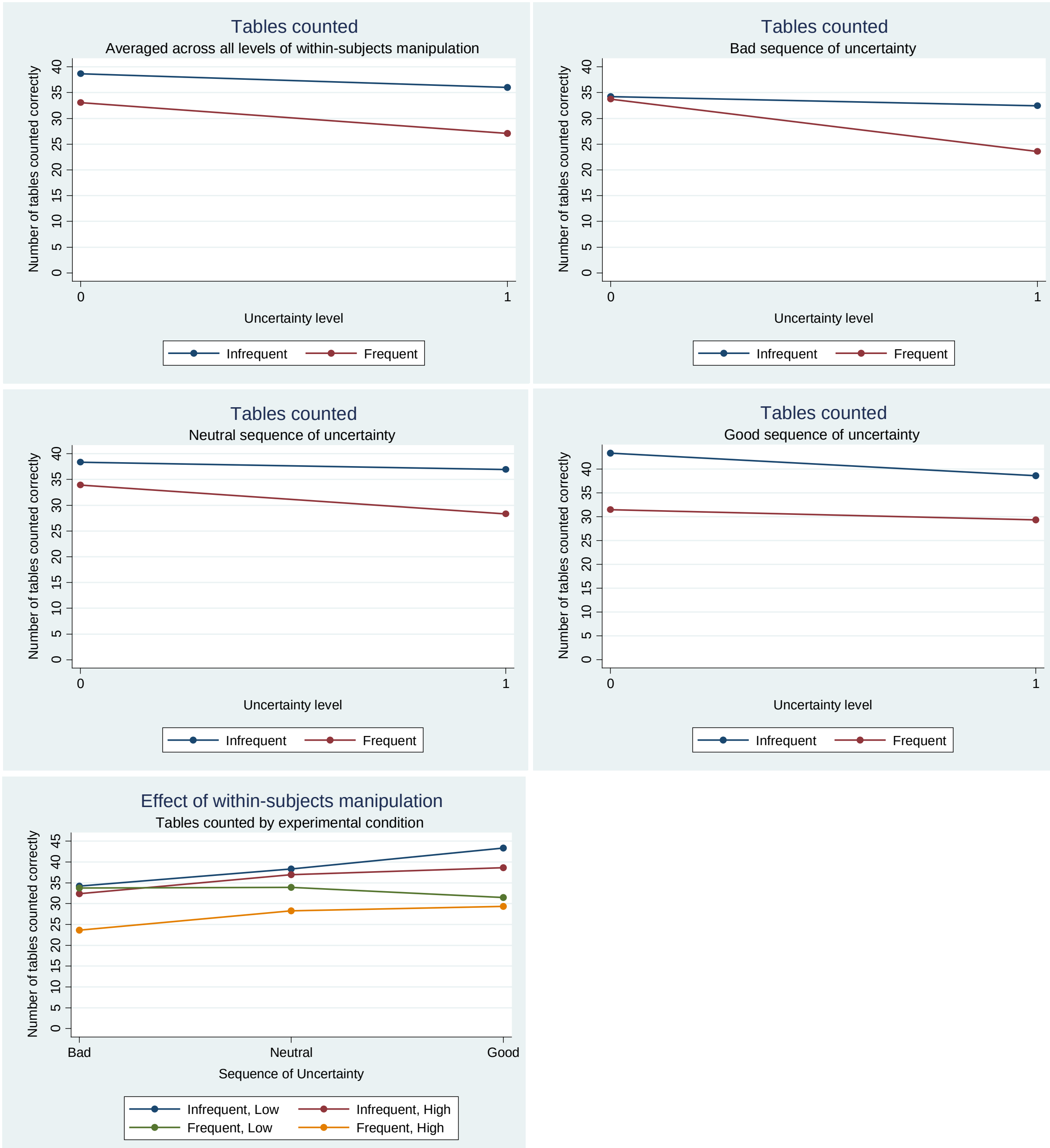


Figure 2-11: Participants counted the number of 0's in tables of 1's and 0's in two rounds, a practice round and a main round. The practice round consisted of one 10-minute period, and the main round consisted of three 10-minute periods. Participants earned points for successfully counting tables. In the practice round, they earned ten points for every correctly counted table. In the main round, participants were awarded 8, 10, or 12 points per table in the *low uncertainty* condition, and 2, 10, or 18 points in the *high uncertainty* condition. For each period of the main round, participants were given a point target and could increase their pay for each period in which they earned their target. In one of the three periods of the main round, participants received mostly low point awards (i.e. 2 or 8 points) early in the period and mostly high point awards (i.e. 12 or 18 points) late in the period; this is referred to as the *bad early sequence* of realizations of uncertainty since bad outcomes resulted early. The *good early sequence* is the opposite – good outcomes occurred early in the period. In the *neutral sequence*, point awards for tables cycled through all

possible values. Each participant experienced each sequence once. *Frequent (infrequent)* refers to the frequency at which participants were given feedback on their progress towards their target. Participants in the *frequent* (*infrequent*) experimental condition received information about points accumulated since the last feedback report every 1 (4) minute(s). They were given the option to stop working towards their target at each feedback report and instead browse the web.

The vertical axis, number of tables counted correctly, refers to the number of tables for which participants received points. The top left figure shows the average of this measure across all three levels of the sequence manipulation. The top right, middle left, and middle right figures show tables counted in the bad, good, and neutral sequences of uncertainty, respectively. The bottom left figure shows the number of tables counted by sequence, with one line for each level of the frequency and uncertainty manipulations.

Table 2-2: Cell means of dependent variables

Participants counted the number of 0's in tables of 1's and 0's in two rounds, a practice round and a main round. The practice round consisted of one 10-minute period, and the main round consisted of three 10-minute periods. Participants earned points for successfully counting tables. In the practice round, they earned ten points for every correctly counted table. In the main round, participants were awarded 8, 10, or 12 points per table in the low uncertainty condition, and 2, 10, or 18 points in the high uncertainty condition. For each period of the main round, participants were given a point target and could increase their pay for each period in which they earned their target. In one of the three periods of the main round, participants received mostly low point awards (i.e. 2 or 8 points) early in the period and mostly high point awards (i.e. 12 or 18 points) late in the period; this is referred to as the bad early sequence of realizations of uncertainty since bad outcomes resulted early. The good early sequence is the opposite – good outcomes occurred early in the period. In the neutral sequence, point awards for tables cycled through all possible values. Each participant experienced each sequence once. Frequent (infrequent) refers to the frequency at which participants were given feedback on their progress towards their target. Participants in the frequent (infrequent) experimental condition received information about points accumulated since the last feedback report every 1 (4) minute(s). They were given the option to stop working towards their target at each feedback report and instead browse the web.

Tables counted refers to the number of tables for which participants received points (i.e. the number of correct responses submitted). Proportion of participants who abandoned refers to the proportion of participants in each experimental condition who exercised the option to abandon their target before having achieved it.

This table shows performance in each of the experimental conditions. In each cell, the first number (**in bold**) is the dependent measure averaged over all three sequences of uncertainty. The second, third, and fourth numbers are the values for the bad early, neutral, and good early sequences, respectively.

	Tables counted				Proportion of participants who abandoned			
Feedback frequency	N	Low uncertainty	High uncertainty	Diff	N	Low uncertainty	High uncertainty	Diff
Infrequent	39	**38.6**	**36**	2.6	39	**21.7%**	**24.6%**	-2.9
		34.2	32.4			35.0	42.1	
		38.4	36.9			25.0	26.3	
		43.4	38.6			5.0	5.3	
Frequent	39	**33.0**	**27.1**	5.9	39	**31.6%**	**40.0%**	-8.4
		33.7	23.6			31.6	55.0	
		33.9	28.3			26.3	35.0	
		31.5	29.4			36.8	30.0	
Difference		5.6	8.9			-9.9	-15.4	

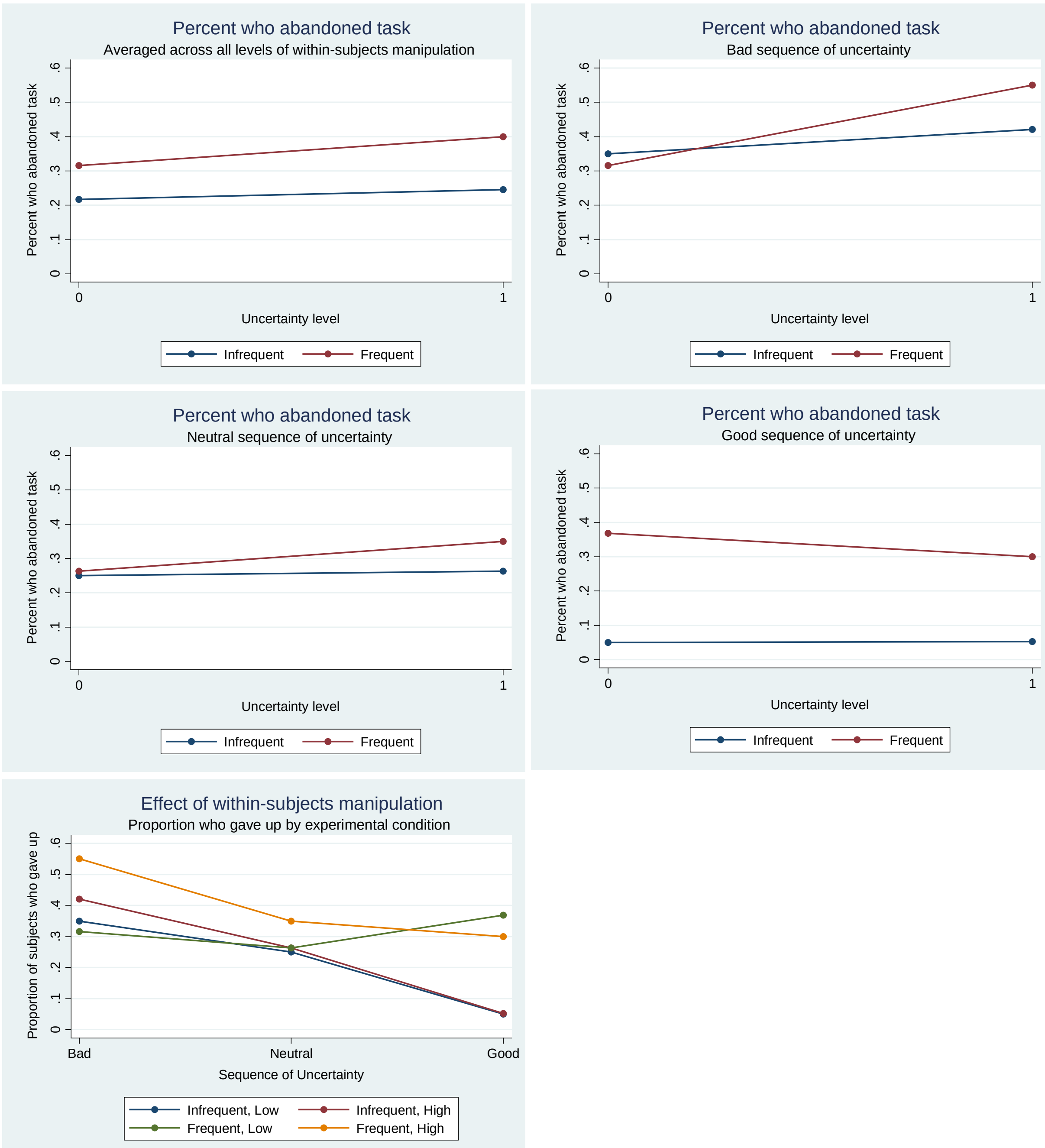


Figure 2-12: Participants counted the number of 0's in tables of 1's and 0's in two rounds, a practice round and a main round. The practice round consisted of one 10-minute period, and the main round consisted of three 10-minute periods. Participants earned points for successfully counting tables. In the practice round, they earned ten points for every correctly counted table. In the main round, participants were awarded 8, 10, or 12 points per table in the low uncertainty condition, and 2, 10, or 18 points in the high uncertainty condition. For each period of the main round, participants were given a point target and could increase their pay for each period in which they earned their target. In one of the three periods of the main round, participants received mostly low point awards (i.e. 2 or 8 points) early in the period and mostly high point awards (i.e. 12 or 18 points) late in the period; this is referred to as the bad early sequence of realizations of uncertainty since bad outcomes resulted early. The good early sequence is the opposite – good outcomes occurred early in the period. In the neutral sequence,

point awards for tables cycled through all possible values. Each participant experienced each sequence once. Frequent (infrequent) refers to the frequency at which participants were given feedback on their progress towards their target. Participants in the frequent (infrequent) experimental condition received information about points accumulated since the last feedback report every 1 (4) minute(s). They were given the option to stop working towards their target at each feedback report and instead browse the web.

The vertical axis, percent who abandoned, refers to the proportion of participants who abandoned their target before achieving it. The top left figure shows the average of this measure over all three levels of the sequence manipulation. The top right, middle left, and middle right figures show the proportion who abandoned in the bad early, good early, and neutral sequences of uncertainty, respectively. The bottom left figure shows the proportion of participants who abandoned by sequence, with one line for each level of the frequency and uncertainty manipulations.

2-5.2 Test of hypotheses

2-5.2.1 Changes in rate of effort after a feedback report

To test whether participants altered their rate of effort exertion following a feedback report indicating greater or lesser performance than expected, I ran a fixed-effects panel regression of number of tables counted in the one minute following a feedback report on unexpected progress. Expected progress at any time is the average number of tables that must be counted per minute to achieve the target. Unexpected progress is actual number of tables counted minus the expected number of tables required since the last feedback report. A fixed-effects model is appropriate here because one of the independent variables, unexpected progress, is likely to be correlated with a subject-specific effect, table counting speed (Cameron and Trivedi 2009). Hypothesis 1 predicts that participants will alter their rate of effort exertion upon receiving a feedback report indicating that performance deviated from expected progress. Thus, a statistically significant coefficient on the unexpected progress variable indicates support for hypothesis 1. I also ran a fixed-effects panel regression of the number of tables counted in the one minute following a feedback report on the *sign* of expected progress. Again, a statistically significant coefficient on sign of unexpected progress will indicate support for hypothesis 1.

The results of the two panel regressions are shown in Table 2-3. The coefficient on unexpected progress in model 1 is positive but not statistically significant, while the coefficient on the sign of unexpected progress in model 2 is *positive* and statistically significant. This result provides marginal support for hypothesis 1. Participants weakly increased (decreased) effort following a reporting indicating greater (lower) progress than expected[25]. Untabulated debriefing responses indicate that participants felt disappointed upon receiving a progress report indicating lower performance than expected ($p < 0.0001$), and that they felt elated upon receiving a progress report indicating greater performance than expected ($p < 0.0001$). Thus it appears that participants experienced affect upon receiving performance feedback, consistent with Carver and Scheier (1998) prediction. However, upon receiving that feedback, they behaved opposite to the predictions of (Carver and Scheier 1998), but consistently with the findings in Erez and Isen (2002). Effort increased after receiving good news and decreased after receiving bad news. Neither of the between-subjects dummy variables, nor their interaction, are significant, consistent with a psychological effect that is not predicted to vary between experimental conditions.

Table 2-3: Panel regression with fixed effects of tables counted in minute T given unexpected performance provided in a feedback report at the end of minute T-1

The table shows the results of two fixed-effects panel regressions of the number of tables of 0's and 1's correctly counted in minute T given unexpected progress reported in feedback at the end of minute T-1. Expected progress at any time is the average number of tables that must be counted per minute to achieve the target. Unexpected progress is actual number of tables counted minus the expected number of tables required since the last feedback report. Thus, unexpected progress was positive if actual progress exceeded expected progress. The Frequency dummy was 1 if participants were in the frequent feedback experimental condition, where feedback was provided every 1 minute, and 0 if participants were provided feedback every 4 minutes. The Uncertainty dummy was 1 if participants point awards per table were drawn from 2, 10, or 18 points, and was 0 if participants point awards per table were drawn from 8, 10, or 12 points per table.

Model (1) is a fixed-effects panel regression of tables counted in minute T on the unexpected progress provided in the feedback report at the end of minute T-1. Model (2) is a fixed-effects panel regression of tables counted in minute T on the *sign* of unexpected progress provided in the feedback report at the end of minute T-1. A '*', '**', or '***' after a coefficient indicates significance at the 10%, 5%, or 1% level, respectively.

[25] Another possibility is that, since the average participant can count 4 tables per minute, and feedback reports came every minute for some participants, it was not possible for them to measurably increase their counting rate. For example, if someone now counts 4.25 tables per minute, the data would not pick that up on a per minute basis.

DV: Effort in period T	(1) Model	(2) Model
Unexpected progress	0.0005	
	(0.0009)	
Sign of unexpected progress		0.21*
		(0.11)
Constant	4.33***	4.21***
	(0.07)	(0.12)
Observations (groups)	308 (73)	308 (73)
R^2 within	0.0011	0.0083
R^2 between	0.0901	0.2516
R^2 overall	0.0341	0.0760

2-5.2.2 Variation in effort due to feedback frequency and uncertainty level

Hypothesis 2a predicts that effort will be lowest when feedback frequency is high and uncertainty is high. People will incorrectly infer from recent information that they are either too far behind schedule, or too far ahead of schedule to achieve their target. As a result, they will decrease effort. To test this hypothesis, I used a planned contrast to test whether the mean of effort in the frequent feedback, high uncertainty condition was lower than the mean of the other three cells. Specifically, I tested whether:

$$\mu_{FH} < \frac{1}{3}(\mu_{FL} + \mu_{IH} + \mu_{IL})$$

Where the subscripts 'F' and 'I' denote the frequent and infrequent feedback, respectively, and the subscripts 'H' and 'L' denote the high and low uncertainty conditions, respectively.

The contrast was a post-estimation test that followed a repeated-measures ANOVA. Repeated measures ANOVA is appropriate because I obtained three measurements of effort for each participant, one for each sequence of uncertainty. Participants were randomly assigned to frequency and uncertainty conditions (between subjects), but every subject experienced one period each of the bad early, good early, and neutral sequences of uncertainty.

Panels A and B of Table 2-4 show the repeated measures ANOVA and contrast test, respectively for hypothesis 1. The contrast test reported in panel B is highly significant (p < 0.0001, one-tailed[26]), providing strong support for hypothesis 2a, and indicating that participants counted approximately nine fewer tables per 10-minute period in the frequent feedback, high uncertainty condition than they did in the other three conditions. The repeated measures ANOVA reported in Panel A of Table 2-4 shows a significant main effect of feedback frequency on effort ($p = 0.0141$). This is consistent with Lam et al. (2011), which implies that effort would be lower in the frequent feedback conditions because processing feedback is cognitively taxing and involves more interactions. The ANOVA also shows an insignificant interaction between frequency and uncertainty ($p = 0.5674$). However, because I am testing for a very specific pattern of results, that effort in one cell is lower than the other three, a planned contrast is a more appropriate statistical test, (Buckless and Ravenscroft 1990). The interaction would be more appropriate for testing a disordinal interaction. The repeated-measures ANOVA also shows a significant main effect of sequence of uncertainty; this will be examined in greater detail below.

Table 2-4: Repeated-measures ANOVA and contrast test of effort dependent variable

Participants counted the number of 0's in tables of 1's and 0's in two rounds, a practice round and a main round. The practice round consisted of one 10-minute period, and the main round consisted of three 10-minute periods. Participants earned points for successfully counting tables. In the practice round, they earned ten points for every correctly counted table. In the main round, participants were awarded 8, 10, or 12 points per table in the low uncertainty condition, and 2, 10, or 18 points in the high uncertainty condition. For each period of the main round, participants were given a point target and could increase their pay for each period in which they earned their target. In one of the three periods of the main round, participants received mostly low point awards (i.e. 2 or 8 points) early in the period and mostly high point awards (i.e. 12 or 18 points) late in the period; this is referred to as the bad early sequence of realizations of uncertainty since bad outcomes resulted early. The good early sequence is the opposite – good outcomes occurred early in the period. In the neutral sequence, point awards for tables cycled through all possible values. Each participant experienced each sequence once. Frequent (infrequent) refers to the frequency at which participants were given feedback on their progress towards their target. Participants in the frequent (infrequent) experimental condition received information about points accumulated since the last feedback report every 1 (4) minute(s). They were given the option to stop working towards their target at each feedback report and instead browse the web.

---

[26] MacDonald (2011) states that the F-distribution with 1 numerator degree of freedom and d denominator degrees of freedom is equivalent to the square of the Student t-distribution with d degrees of freedom. Therefore, "*[a]s long as the F test has 1 numerator degree of freedom, the square root of the F statistic is the absolute value of the t statistic for the one-sided test.*" Since all contrast tests reported in this paper are Wald tests with 1 numerator degree of freedom, this method is appropriate, and is equivalent to dividing the two-tailed p-value in half.

Panel A shows a repeated measures ANOVA of the effort dependent measure on the two between-subjects manipulations, feedback frequency and uncertainty, as well as the within-subjects manipulation, the sequence of uncertainty. Effort is measured by the number of tables for which participants received points in each 10-minute period. Thus, there are three measurements per subject.

Panel B shows a planned contrast that tests whether the mean number of tables counted in the Frequent feedback, High uncertainty condition ("FH") is lower than the average number of tables counted in the other three conditions: Frequent feedback, Low uncertainty ("FL"); Infrequent feedback, High uncertainty ("IH"); Infrequent feedback, Low uncertainty ("IL").

Panel A: Repeated-measures ANOVA of tables counted on feedback frequency, uncertainty, and sequence of realizations of uncertainty

Num. of obs	234	R^2	0.7131
Root MSE	10.7069	Adjusted R^2	0.5483

Source	Partial SS	df	MS	F	Prob > F
Between subjects:	4413.55	3	1471.18	3.02	0.0352
Frequency	3079.21	1	3079.21	6.32	0.0141
Uncertainty	1077.21	1	1077.21	2.21	0.1414
Freq. x Uncertainty	160.94	1	160.94	0.33	0.5674
Subject x Freq. x Unc.	36081.78	74	487.59		
Within subjects:	1676.39	8	209.55	1.83	0.0761
Sequence	920.16	2	460.08	4.01	0.0201
Seq. x Frequency	358.00	2	179.00	1.56	0.2133
Seq. x Uncertainty	81.13	2	40.56	0.35	0.7026
Seq. x Freq. x Unc.	297.79	2	148.89	1.30	0.2759
Residual	16966.28	148	114.64		
Total	59138	233	253.81		

Panel B: Contrast test of effort in frequent feedback, high uncertainty cell against other feedback and uncertainty conditions

	Effect size	F	p-value (1-tailed)
Contrast:			
Effort in frequent feedback, high uncertainty < Average effort in other three between-subjects conditions $\mu_{FH} < \frac{1}{3}(\mu_{FL} + \mu_{IH} + \mu_{IL})$	8.81	10.13	0.0000

2-5.3 Effect of feedback frequency and uncertainty on likelihood of giving up

Hypothesis 2b builds on hypothesis 2a. H2a predicts that effort will be lowest when feedback frequency is high and uncertainty is high. H2b predicts that the driving factor behind H2a is participants' decision to give up on their targets and instead browse the web in the *Leisure Zone*. Participants will incorrectly infer from recent information that they are either too far behind schedule, or too far ahead of schedule to achieve their target. At some point, they will be so far behind that they will assess their chances of achieving their goals and decide that they are unlikely to achieve their target and earn the target bonus, and are therefore better off not working.

To test this hypothesis, I used a planned contrast to test whether the proportion of giving up in the frequent feedback, high uncertainty condition was higher than the mean of the other three cells. Specifically, I tested whether:

$$\mu_{FH} > \frac{1}{3}(\mu_{FL} + \mu_{IH} + \mu_{IL})$$

Where the subscripts 'F' and 'I' denote the frequent and infrequent feedback, respectively, and the subscripts 'H' and 'L' denote the high and low uncertainty conditions, respectively.

The contrast was a post-estimation test that followed a logistic random-effects panel regression with random effects. A random-effects panel regression is appropriate because I obtained three measurements of giving up for each participant, one for each sequence of uncertainty. Participants were randomly assigned to frequency and uncertainty conditions (between subjects), but every subject experienced one period each of the bad early, good early, and neutral sequences of uncertainty. A logistic regression is appropriate because the dependent measure is a binary decision. Finally, a random effects model is appropriate because participants were randomly assigned to treatments and therefore the right-hand side variables of the regression are uncorrelated with the individual-specific effects (Cameron and Trivedi 2009).

Panels A and B of Table 2-5 show the logistic panel regression and contrast test, respectively for hypothesis 1. The contrast test reported in panel B is highly significant (p = 0.0333, one-tailed[27]), providing strong support for hypothesis 2, and indicating that participants were approximately 1.27 times more likely to give up in the frequent feedback, high uncertainty condition than they were in the other three conditions. The logistic regression in Panel A of Table 2-5 also shows a significant main effect of the good early sequence of uncertainty (p = 0.016). The coefficient is negative and of large magnitude, indicating that participants who observed good outcomes early in the period were much less likely to give up and enter the *Leisure Zone*. This finding is consistent with the notion that feedback indicating significant goal progress leads to positive affect. However, it is not consistent with the theory that predicts slacking off and ultimately quitting as a result of falling behind. It appears more likely that participants were energized and motivated by early positive results and this carried forward through the period. Untabulated results from debriefing questions demonstrate that participants did feel good after receiving good news; there were no differences between experimental conditions. This effect, reduced likelihood of giving up with early good outcomes, was entirely reversed when feedback frequency was high, as indicated by the interaction between feedback frequency and the good early sequence of uncertainty (p = 0.024). Thus, the frequent reassessments due to frequent feedback counteracts the positive affect from good news, leading participants to quit when not observing sufficiently high progress later in the period when the noise reverted.

Table 2-5: Logistic random-effects panel regression and contrast test of abandonment dependent variable

Participants counted the number of 0's in tables of 1's and 0's in two rounds, a practice round and a main round. The practice round consisted of one 10-minute period, and the main round consisted of three 10-minute periods. Participants earned points for successfully counting tables. In the practice round, they earned ten points for every correctly counted table. In the main round, participants were awarded 8, 10, or 12 points per table in the low uncertainty condition, and 2, 10, or 18 points in the high uncertainty condition. For each period of the main round, participants were given a point target and could increase their pay for each period in which they earned their target. In one of the three periods of the main round, participants received mostly low point awards (i.e. 2 or 8 points) early in the period and mostly high point awards (i.e. 12 or 18 points) late in the period; this is referred to as the bad early sequence of realizations of uncertainty since bad outcomes resulted early. The good early sequence is the opposite – good outcomes occurred early in the

[27] Ibid.

period. In the neutral sequence, point awards for tables cycled through all possible values. Each participant experienced each sequence once. Frequent (infrequent) refers to the frequency at which participants were given feedback on their progress towards their target. Participants in the frequent (infrequent) experimental condition received information about points accumulated since the last feedback report every 1 (4) minute(s). They were given the option to stop working towards their target at each feedback report and instead browse the web.

Panel A shows a logistic random-effects panel regression with random effects. The dependent variable is the proportion participants in each experimental condition who chose to abandon their target, before achieving it, and browse the web. The independent variables are categorical and represent the experimental conditions, frequency, uncertainty, and sequence, and their interactions. Because each subject experienced each value of the sequence manipulation, there are three measurements per subject.

Panel B shows a planned contrast that tests whether the proportion of participants who abandoned their target in the Frequent feedback, High uncertainty condition ("FH") is higher than the average number of tables counted in the other three conditions: Frequent feedback, Low uncertainty ("FL"); Infrequent feedback, High uncertainty ("IH"); Infrequent feedback, Low uncertainty ("IL"). A '*', '**', or '***' after a coefficient indicates significance at the 10%, 5%, or 1% level, respectively.

Panel A: Logistic panel regression with random effects of proportion of participants who abandoned goal on feedback frequency, uncertainty, and sequence of realizations of uncertainty

	Dependent variable: Entered Leisure Zone
Frequency	-0.25
	(1.09)
Uncertainty	0.50
	(1.06)
Frequency x Uncertainty	1.10
	(1.51)
SEQUENCE	
Neutral	-0.77
	(0.89)
Good early	-3.32***
	(1.37)
FREQUENCY x SEQUENCE	
Frequent x Neutral	0.34
	(1.28)
Frequent x Good early	3.71**
	(1.64)
UNCERTAINTY x SEQUENCE	
High x Neutral	-0.32
	(1.24)
High x Good	-0.27
	(1.87)
FREQ. X UNC. X SEQ.	
Frequent x High x Neutral	-0.54
	(1.75)
Frequent x High x Good	-1.76
	(2.25)

Constant	-1.03
	(0.77)
Wald χ^2_{11}	17.15
Prob > χ^2	0.1036
Num. of obs. (subjects)	234 (78)

Panel B: Contrast test of proportion who abandoned their target in frequent feedback, high uncertainty cell against other feedback and uncertainty conditions

	Effect size	χ^2	p-value (1-tailed)
Contrast:			
Proportion giving up in frequent feedback, high uncertainty > Average proportion in other three between-subjects conditions			
$\mu_{FH} > \frac{1}{3}(\mu_{FL} + \mu_{IH} + \mu_{IL})$	1.27	3.36	0.0333

2-5.4 Effects of sequences of uncertainty realizations on effort provision

Hypothesis 3 predicts how the sequence manipulation should affect the pattern of effort exerted by participants. Specifically, it predicts that when feedback frequency is high and uncertainty is high, effort should be lowest when participants encounter the bad early sequence of uncertainty in which bad outcomes occur early. Effort should be higher for the good s early equence, and highest for the neutral sequence. Effort is predicted to be lower for the good early sequence than the neutral sequence since participants are expected to coast on the early good outcomes and then later realize that their goal is unachievable.

I used a set of planned contrasts to test hypothesis 3. These contrast tests were run as post-estimation tests to the repeated measures ANOVA reported in Table 2-4. The specific contrasts, which apply only to the frequent feedback, high uncertainty ("FH") condition, are:

$$\mu_{FH}^{\text{Bad}} < \mu_{FH}^{Good}$$

$$\mu_{FH}^{\text{Bad}} < \mu_{FH}^{Neutral}$$

$$\mu_{FH}^{\text{Good}} < \mu_{FH}^{Neutral}$$

The pattern of results is illustrated in the last graph of Figure 2-11. The contrasts, as well as a joint test of all three contrasts, are shown in Table 2-6. These tests indicate that participants do indeed exert less effort when the encounter the bad early sequence of uncertainty than when they encounter the good (p = 0.0458). They also exert less effort under the bad early sequence than under the neutral sequence (p = 0.0836). However, effort exerted under the good early and neutral sequences is statistically indistinguishable (p = 0.3785). Thus, it appears that participants did exert less effort under the bad early sequence of uncertainty, but the predicted effect of slacking off and later giving up under the good early sequence of uncertainty does not appear to have occurred. Another possibility is that, contrary to the slacking off hypothesis, participants who received early positive feedback were elated and worked harder, thus behaving contrary to the prediction of hypothesis 3. The data in Table 2-3 support this idea, and that is likely why only marginal support for hypothesis 3 is found. A joint test of the three planned contrasts composing hypothesis 3 is barely significant at the 10% level (p = 0.0992).

Table 2-6: Contrast tests to establish the effect of sequence of uncertainty realizations on effort when frequent feedback provided in highly uncertain environments

The table shows the planned contrasts needed to test hypothesis 3, which predicts a specific pattern of effort that results from the sequence manipulation within the frequent feedback, high uncertainty condition. Participants counted the number of 0's in tables of 1's and 0's for three 10-minute periods. In each period, a different sequence of point awards was presented to participants. In one of the three periods of the main round, participants received mostly low point awards (i.e. 2 or 8 points) early in the period and mostly high point awards (i.e. 12 or 18 points) late in the period; this is referred to as the *Bad early* sequence of realizations of uncertainty since bad outcomes resulted early. The *Good early* sequence is the opposite – good outcomes occurred early in the period. In the *Neutral* sequence, point awards for tables cycled through all possible values. Each participant experienced each sequence once. The order of sequences was randomized across periods for each participant. The planned contrasts test whether effort, as measured by the number of tables counted in each 10-minute period, differs systematically across periods and follow the pattern predicted in Hypothesis 3. All contrasts are one-tailed.

	df	F	P > F
Bad early < Good early	1	2.88	0.0458
Bad early < Neutral	1	1.93	0.0836
Good early < Neutral	1	0.10	0.3785
Joint test	2	1.64	0.0992

2-5.5 Supplementary analysis

I test for a recency effect: do participants overreact to recent feedback and underreact to aggregate feedback, as Lurie and Swaminathan (2009) find? To test for a recency effect, I tested whether participants quit when their economically optimal choice would have been to continue working on their target, or whether they continued working when they should have quit. To complete this analysis, I assumed that when given a feedback report, participants reassessed their chances of achieving the goal (untabulated debriefing analysis of the question, "*After receiving feedback, I reassessed my chance of hitting the point target and earning $5.*" A response of +3 (-3) indicated strong (dis-)agreement, and zero indicated a neutral response. A t-test of the mean response against the null hypotheses of zero indicates that participants did indeed reassess their chances, $p < 0.0001$).

First, I established a benchmark for the economically optimal decision to quit. At each feedback report, I used information about the rate at which participants counted tables during the experiment to predict the number of tables they could count during the remainder of the period[28]. In other words, if they counted at a slightly accelerated rate, what was their chance of hitting their target, given the time remaining and their current point total? To compute this chance, I simulated a discrete distribution of point outcomes, assuming each extra table counted would be worth 2, 10, or 18 points (8, 10, or 12 points) in the high (low) uncertainty condition. Multiplying this chance by the $5 value of the target yielded an expected value of the target.

I compared the expected value of the target at each point in time to the amount of money available if the participant entered the *Leisure Zone* at that time. Assuming risk neutrality in small stakes, I classified participants who quit working on tables and entered the *Leisure Zone* when the expected value of the target exceeded the certain amount of money for entering the *Leisure Zone* as early quitters. I classified those who kept working when the amount of money available in the *Leisure Zone* exceeded the expected value of the target as stubborn. If a recency effect was present, early quitters and stubborn participants should appear more often in the frequent feedback

[28] Specifically, I looked at the number of tables that participant counted during each 3-minute window of the experiment, including the practice round. This gave a set of counting rates for each participant. I then took the 75th percentile rate and assumed that, if they exerted effort, they should be able to consistently count at this 75th percentile rate.

condition. The mean proportions of early quitters and stubborn participants are shown in Table 2-7. The table reveals no real difference in early quitters by experimental condition. An untabulated logistic random-effects panel regression appears to confirm this claim, as the table shows differences in stubborn participants across experimental conditions. However, an untabulated logistic random-effects panel regression reveals no significant difference in stubborn participants by experimental condition.

Table 2-7: Cell means of early quitters and stubborn participants

At each feedback report, I used information about the rate at which participants counted tables during the experiment to predict the number of tables they could count during the remainder of the period. Given this assumed counting rate, and participants' point total at that time, I simulated a discrete distribution of point outcomes. Using this information, I could computed the chance that the participant could achieve their target. Multiplying this chance by the $5 value of the target yielded an expected value of the target. I compared this expected value to the amount of money available if the participant entered the Leisure Zone at that time. I classified participants who quit working on tables and entered the Leisure Zone when the expected value of the target exceeded the certain amount of money for entering the Leisure Zone as early quitters. I classified those who kept working when the amount of money available in the Leisure Zone exceeded the expected value of the target as stubborn. This table shows the proportions of early quitters and stubborn participants by experimental condition.

	Proportion of early quitters		Proportion of stubborn participants	
Feedback frequency	Low uncertainty	High uncertainty	Low uncertainty	High uncertainty
Infrequent	**2.8%**	**2.8%**	**12.8%**	**23.6%**
Frequent	**1.5%**	**2.2%**	**28.3%**	**21.5%**

2-6. Discussion and conclusion

This paper presented the results of an experiment designed to find evidence of a dark side of feedback. It hypothesized and tested the notion that when workers are provided with frequent feedback about their progress towards a goal, and the underlying task is noisy, workers are more likely to abandon their goals if early feedback indicates lack of progress towards the goal. This paper found evidence consistent with this hypothesis. As predicted, the evidence was strongest when early realizations of uncertainty were negative.

This study has implications for performance measurement in accounting settings. Accountants often aggregate data and compile them into performance reports. An open-ended question is how frequently this should be done. Common wisdom is that more information is better, and that more frequent and therefore timely information is better. This study suggests a boundary condition – if feedback is provided too frequently, workers with goals might actually experience reduced motivation. Worker motivation may be higher if information is aggregated over longer horizons since the longer horizons allow noise to cancel itself out to a greater degree. This intuition is consistent with prior studies that conclude that information aggregation, or bunching, might be optimal in some settings (e.g. Dye 2010; Arya et al. 2004).

The results in this study are subject to a number of limitations. The theory in this study relies on a mechanism that involves affect as a mediating factor. However, it is not easy to measure affect during the middle of task such as the one used herein. Affect was measured, very indirectly, using debriefing questions on whether certain outcomes during the experiment made participants feel good or bad. This measurement scheme was not ideal as it relied on participants' memories, which may be biased towards the end of the experiment. Additionally, a robustness check found the opposite of the effect predicted by the cited psychology studies. Thus, the mechanism underlying the results cannot clearly be ascertained.

Another limitation is that the findings presented herein can be explained with either an economic model or with psychology theories. Both theories would predict higher goal abandonment rates in highly uncertain settings where feedback is provided frequently. The main differences in these theories' predictions are that psychology predicts that an affective response drives changes in effort exertion and that a recency bias may influence goal abandonment. The evidence presented in this paper is consistent with some aspects of both theories. Additional experiments may be required to differentiate between economic and psychological effects.

2-7. References

Abeler, J., A. Falk, L. Goette, and D. Huffman. 2011. Reference Points and Effort Provision. *American Economic Review* 101 (2):470-492.

Anand, V. 2012. Performance Standard Horizons, Uncertainty, and Controllability. *Working paper, Cornell University*.

Anderhub, V., S. Gächter, and M. Königstein. 2002. Efficient Contracting and Fair Play in a Simple Principal-Agent Experiment. *Experimental Economics* 5 (1):5-27.

Ariely, D., E. Kamenica, and D. Prelec. 2008. Man's search for meaning: The case of Legos. *Journal of Economic Behavior & Organization* 67 (3-4):671-677.

Arya, A., J. Glover, and P. J. Liang. 2004. Intertemporal aggregation and incentives. *European Accounting Review* 13 (4):643-657.

Bandura, A. 1989. Self-Regulation of Motivation and Action Through Internal Standards and Goal Systems. In *Goal Concepts in Personality and Social Psychology*, edited by L. A. Pervin: Lawrence Erlbaum Associates, 19-86.

Block, M. K., and J. M. Heineke. 1973. Allocation of Effort under Uncertainty - Case of Risk-Averse Behavior. *Journal of Political Economy* 81 (2):376-385.

Buckless, F. A., and S. P. Ravenscroft. 1990. Contrast coding: A refinement of ANOVA in behavioral analysis. *Accounting Review*:933-945.

Cameron, A. C., and P. K. Trivedi. 2009. *Microeconometrics using stata*. Vol. 5: Stata Press College Station, TX.

Carver, C. S., and M. Scheier. 1998. *On the self-regulation of behavior*. Cambridge, UK ; New York, NY, USA: Cambridge University Press.

Datar, S., and M. Gupta. 1994. Aggregation, specification and measurement errors in product costing. *Accounting Review*:567-591.

Dye, R. A. 2010. Disclosure "bunching". *Journal of Accounting Research* 48 (3):489-530.

Erez, A., and A. M. Isen. 2002. The influence of positive affect on the components of expectancy motivation. *Journal of Applied Psychology* 87 (6):1055.

Fehr, E., S. Gachter, and G. Kirchsteiger. 1997. Reciprocity as a contract enforcement device: Experimental evidence. *Econometrica* 65 (4):833-860.

Frederickson, J. R. 1992. Relative Performance Information - the Effects of Common Uncertainty and Contract Type on Agent Effort. *Accounting Review* 67 (4):647-669.

Grunfeld, Y., and Z. Griliches. 1960. Is aggregation necessarily bad? *The Review of Economics and Statistics* 42 (1):1-13.

Heath, C., R. P. Larrick, and G. Wu. 1999. Goals as reference points. *Cognitive Psychology* 38 (1):79-109.

Heckman, J., C. Heinrich, and J. Smith. 1997. Assessing the performance of performance standards in public bureaucracies. *American Economic Review* 87 (2):389-395.

Kachelmeier, S. J., B. E. Reichert, and M. G. Williamson. 2008. Measuring and motivating quantity, creativity, or both. *Journal of Accounting Research* 46 (2):341-373.

Kahneman, D., and A. Tversky. 1979. Prospect Theory - Analysis of Decision under Risk. *Econometrica* 47 (2):263-291.

Kaur, S., M. Kremer, and S. Mullainathan. 2010. Self-Control and the Development of Work Arrangements. *American Economic Review* 100 (2):624-628.

Kluger, A. N., and A. DeNisi. 1996. The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin* 119 (2):254-284.

Kukla, A. 1972. Foundations of an Attributional Theory of Performance. *Psychological Review* 79 (6):454-&.

Lam, C. F., D. S. DeRue, E. P. Karam, and J. R. Hollenbeck. 2011. The impact of feedback frequency on learning and task performance: Challenging the "more is better" assumption. *Organizational Behavior and Human Decision Processes* 116 (2):217-228.

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

Libby, R., R. Bloomfield, and M. W. Nelson. 2002. Experimental research in financial accounting. *Accounting Organizations and Society* 27 (8):777-812.

Lim, S. S., and S. Sunder. 1991. Efficiency of asset valuation rules under price movement and measurement errors. *Accounting Review*:669-693.

Locke, E. A., and G. P. Latham. 2002. Building a practically useful theory of goal setting and task motivation. A 35-year odyssey. *American Psychologist* 57 (9):705-717.

Lurie, N. H., and J. M. Swaminathan. 2009. Is timely information always better? The effect of feedback frequency on decision making. *Organizational Behavior and Human Decision Processes* 108 (2):315-329.

MacDonald, K. 2011. *The results from estimation commands display only two-sided tests for the coefficients. How can I perform a one-sided test?* StataCorp LP, July 2011 [cited November 23 2011]. Available from http://www.stata.com/support/faqs/stat/oneside.html.

Rabin, M. 2002. Inference by believers in the law of small numbers. *The Quarterly Journal of Economics* 117 (3):775-816.

Siebert, W. M. 1986. *Circuits, signals and systems*. Vol. 2. Cambridge, MA: The MIT Press.

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

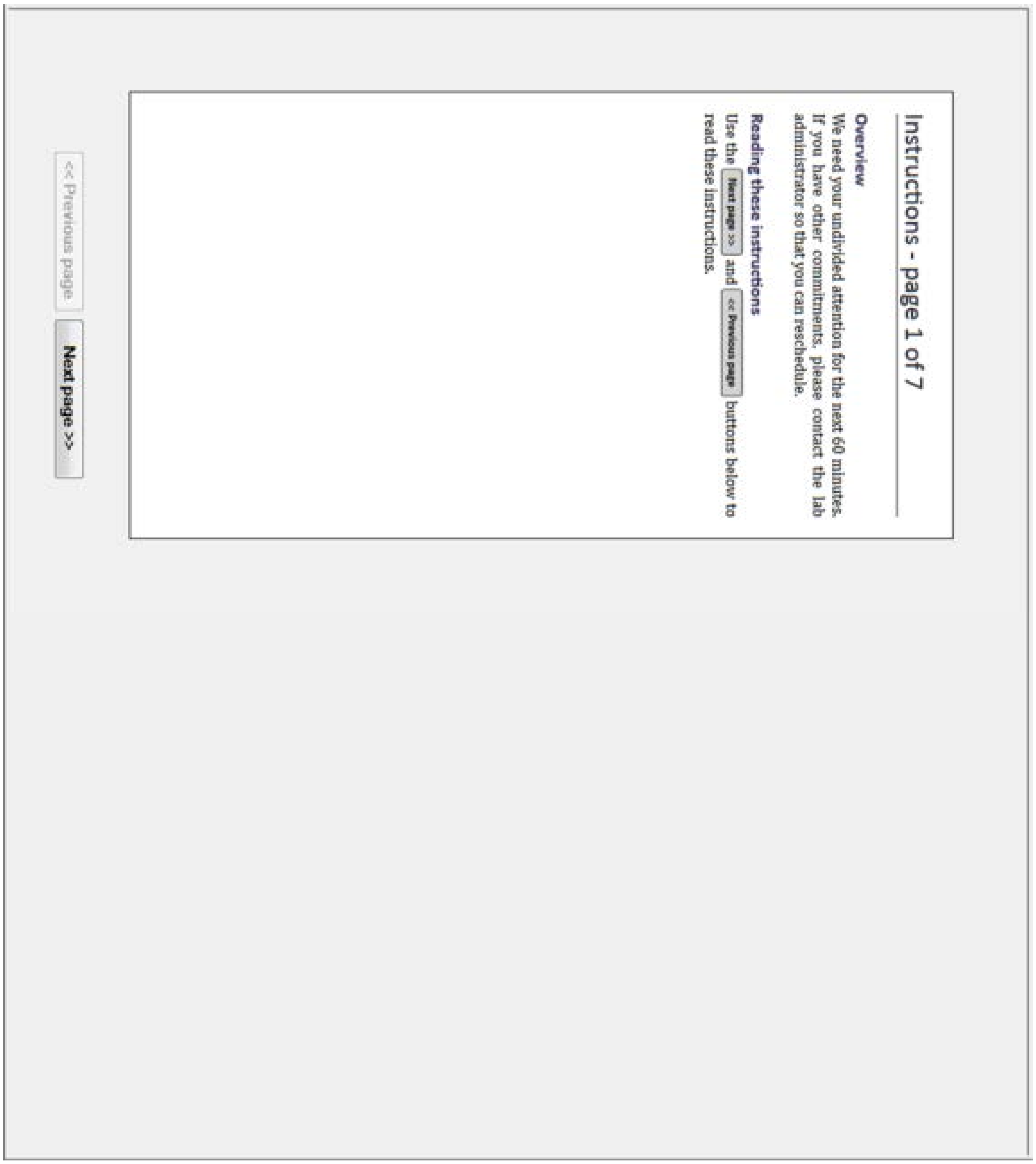
Tversky, A., and D. Kahneman. 1971. Belief in the law of small numbers. *Psychological Bulletin* 76 (2):105.

Vroom, V. H. 1964. Work and motivation.

Weisstein, E. W. 2013. *Weak Law of Large Numbers* 2013 [cited August 16, 2013 2013]. Available from http:/mathworld.wolfram.com/WeakLawofLargeNumbers.html.

2-8. Appendix A – Instructions for practice round

This appendix shows the instructions for the practice round. These instructions were displayed one page at a time using a viewer application written in the C# programming language. The viewer displayed the instructions in a panel on the left, and a sample screen, if available for that page of the instructions, was shown in a separate panel on the right. The panel at right was also used to administer comprehension check questions for the practice round. These questions were presented to participants after they clicked the I'm done with the instructions button on the final page of the instructions.



Instructions - page 2 of 7

Restrictions

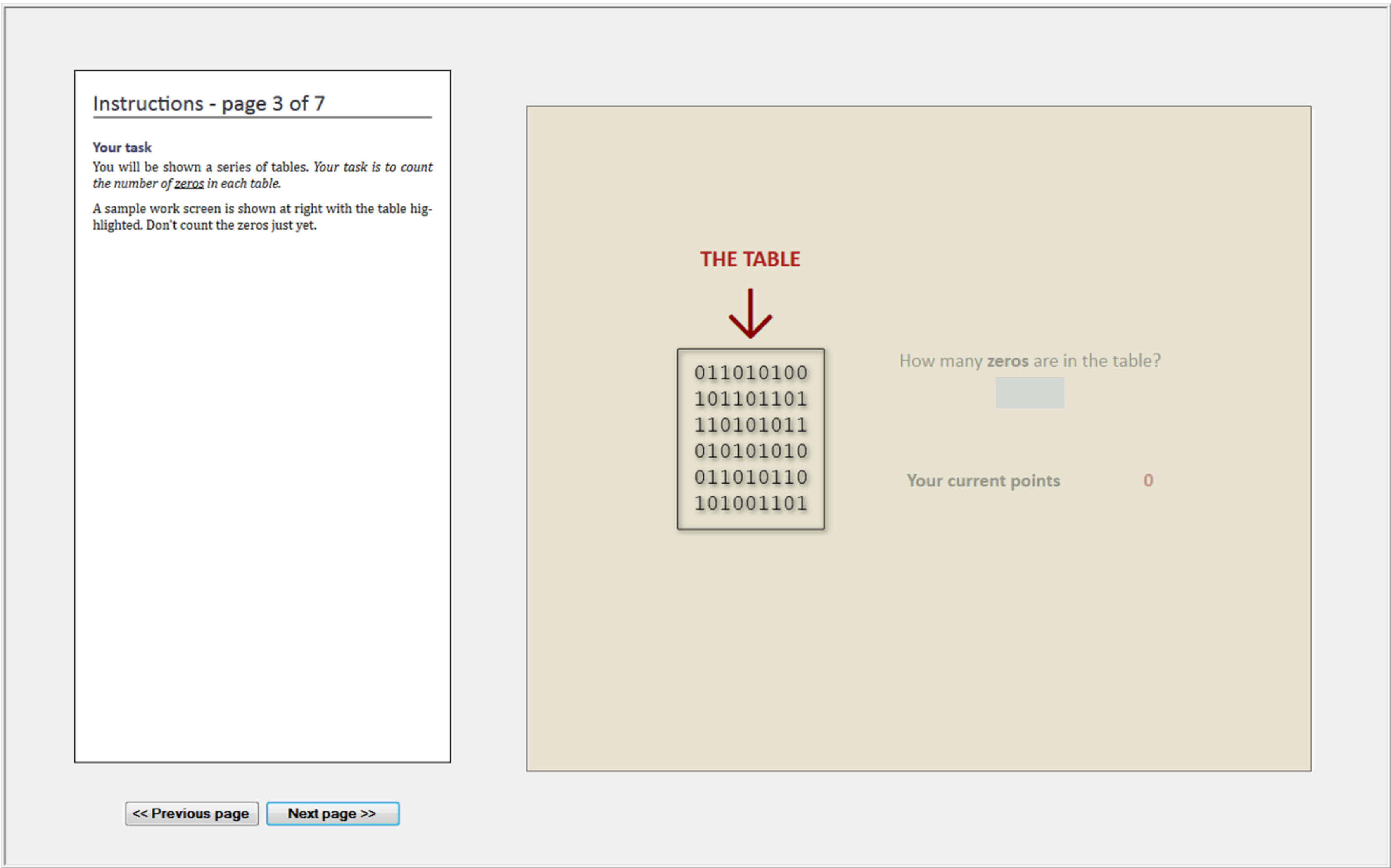
No electronic devices

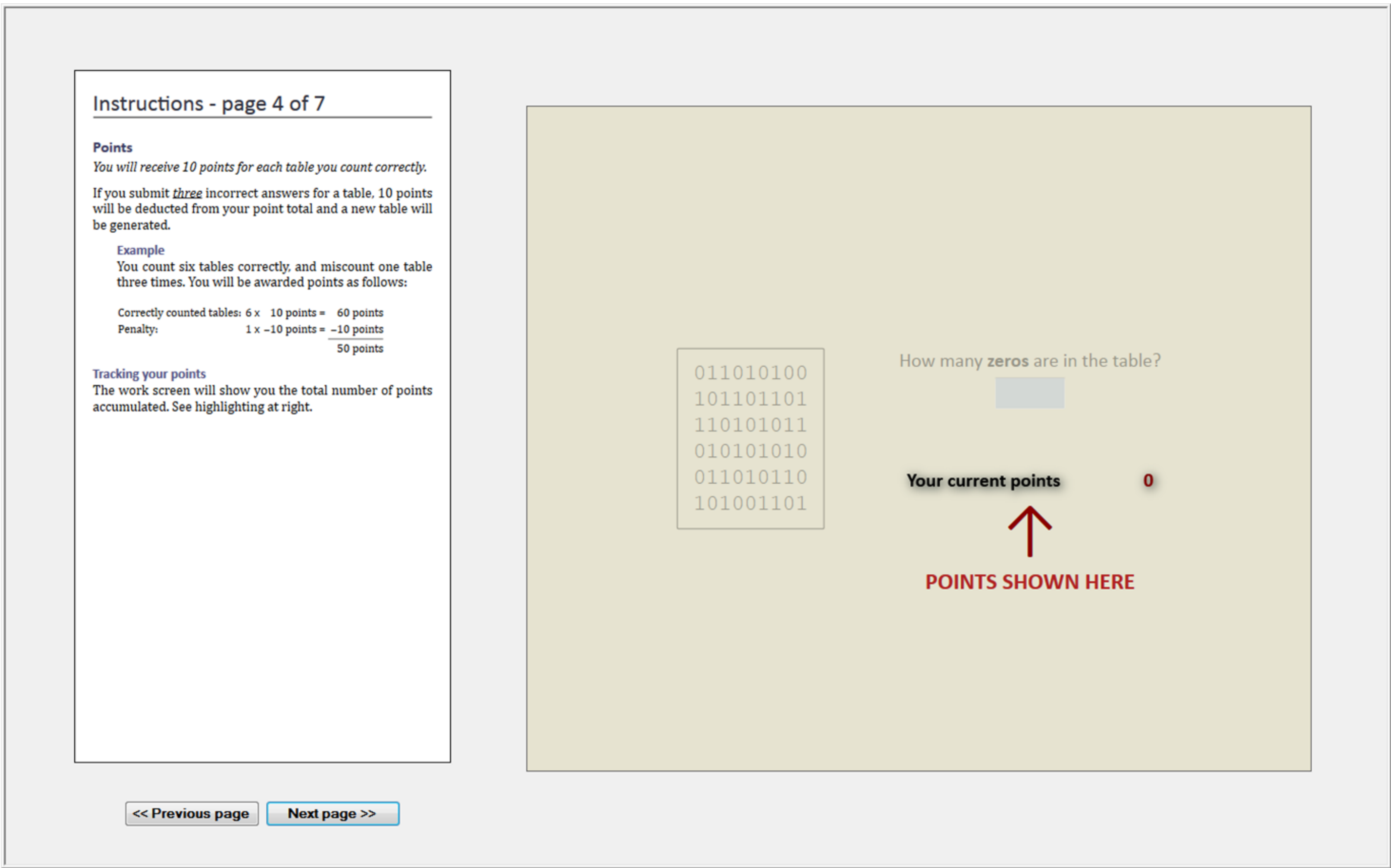
You may not use any electronic devices, such as cell phones, PDA's, iPhones, iPads, and laptops. Please turn off or mute your electronic devices at this time. If you cannot agree to this restriction, then you may not participate.

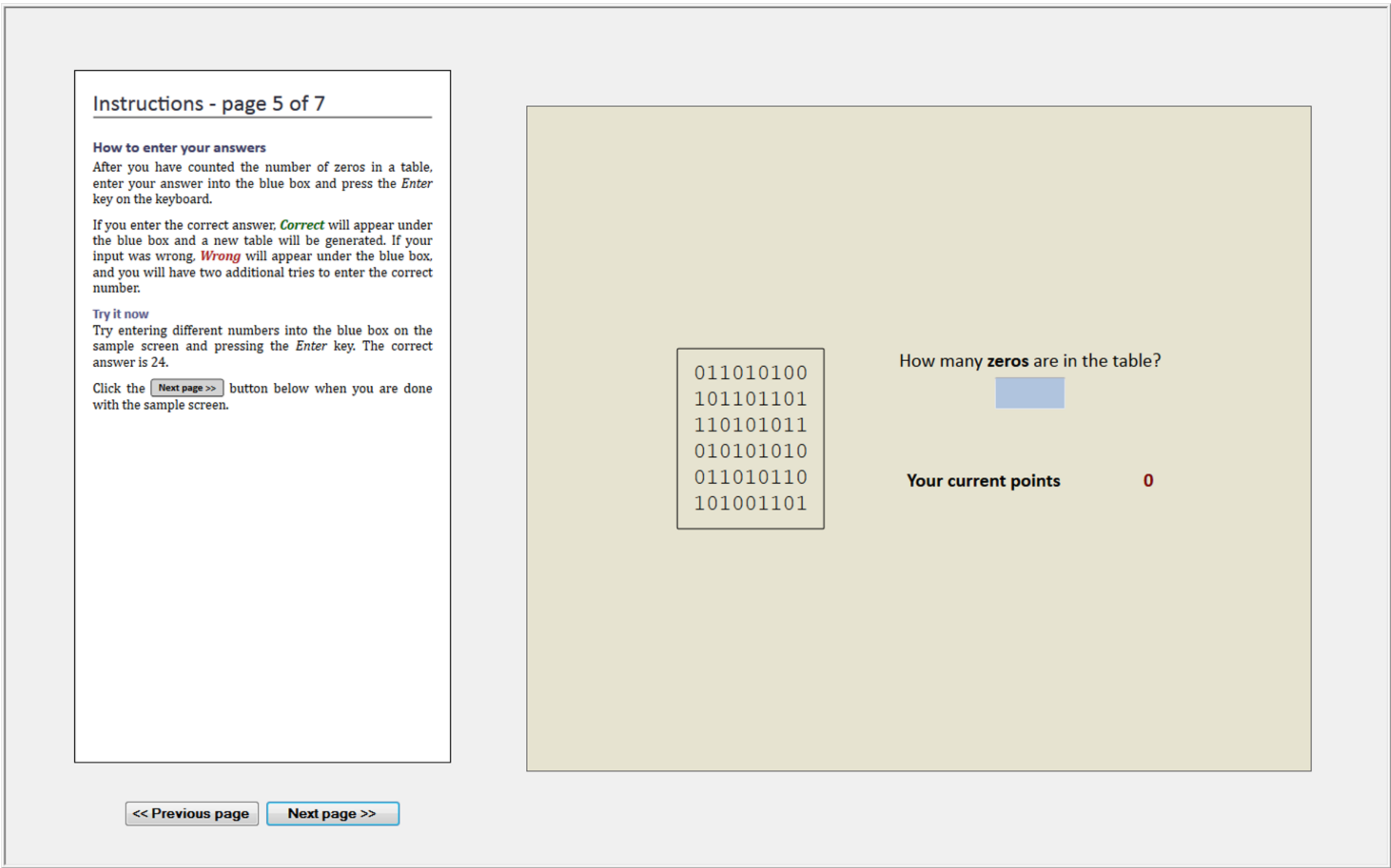
Computer use

You have been assigned a lab computer. You may only use it for the task described below. If you use your assigned computer for any other purpose, you will be asked to leave and will not receive any payment.

<< Previous page | Next page >>


Instructions - page 3 of 7
Your task
You will be shown a series of tables. Your task is to count the number of zeros in each table.
A sample work screen is shown at right with the table highlighted. Don't count the zeros just yet.
THE TABLE
011010100
101101101
110101011
010101010
011010110
101001101
How many zeros are in the table?
Your current points 0
<< Previous page
Next page >>


Instructions - page 4 of 7
Points
You will receive 10 points for each table you count correctly.
If you submit three incorrect answers for a table, 10 points will be deducted from your point total and a new table will be generated.
Example
You count six tables correctly, and miscount one table three times. You will be awarded points as follows:
Correctly counted tables: 6 x 10 points = 60 points
Penalty: 1 x −10 points = −10 points
50 points
Tracking your points
The work screen will show you the total number of points accumulated. See highlighting at right.
<< Previous page
Next page >>
011010100
101101101
110101011
010101010
011010110
101001101
How many zeros are in the table?
Your current points
0
POINTS SHOWN HERE

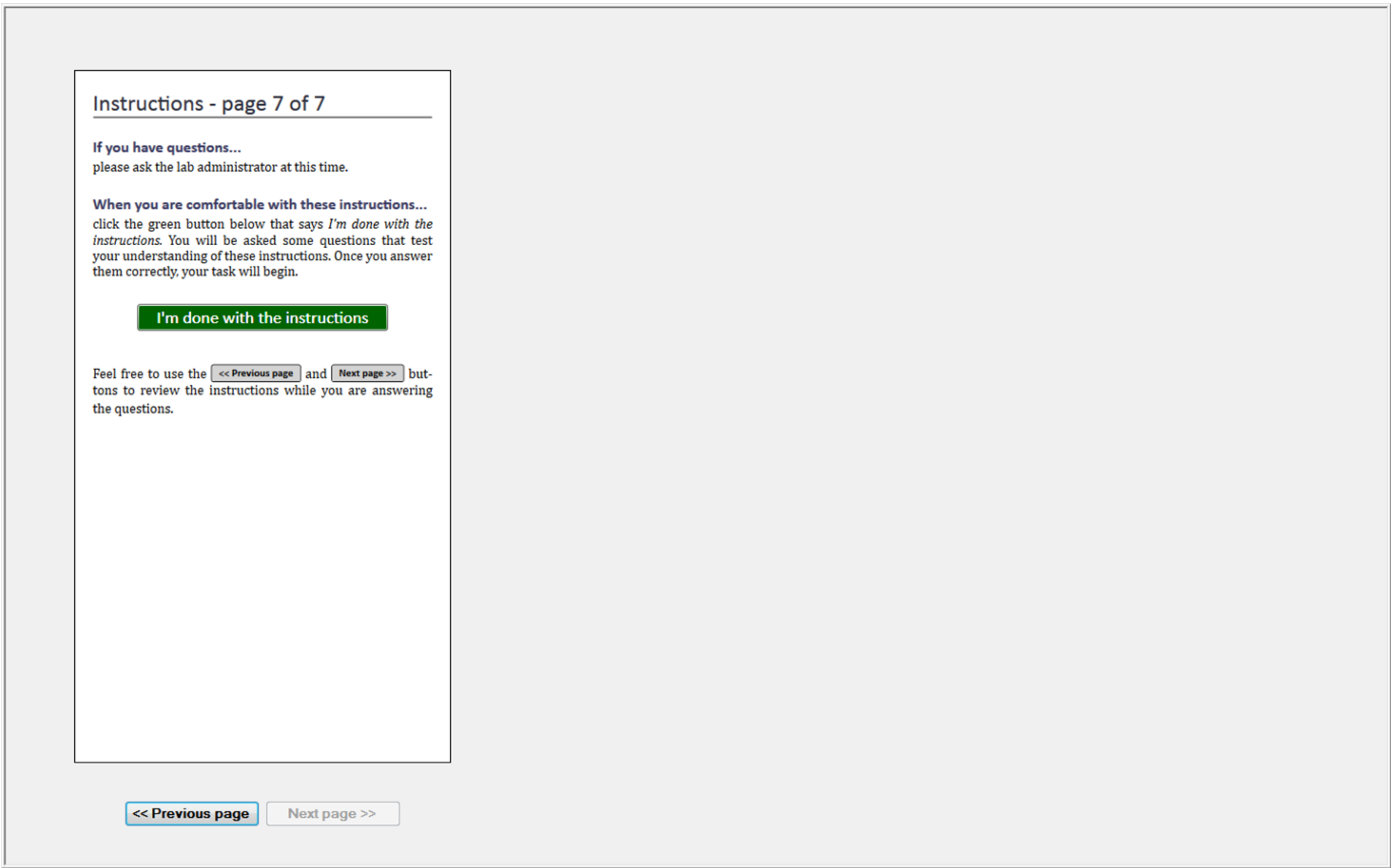

Instructions - page 5 of 7
How to enter your answers
After you have counted the number of zeros in a table, enter your answer into the blue box and press the Enter key on the keyboard.
If you enter the correct answer, Correct will appear under the blue box and a new table will be generated. If your input was wrong, Wrong will appear under the blue box, and you will have two additional tries to enter the correct number.
Try it now
Try entering different numbers into the blue box on the sample screen and pressing the Enter key. The correct answer is 24.
Click the Next page >> button below when you are done with the sample screen.
011010100
101101101
110101011
010101010
011010110
101001101
How many zeros are in the table?
Your current points
0
<< Previous page
Next page >>

Instructions - page 6 of 7

Your pay

- You will be paid a fixed fee of $5.
- In addition, you will be paid 8 cents ($0.08) for every *ten* points that you earn.

<< Previous page | Next page >>


Instructions - page 7 of 7
If you have questions...
please ask the lab administrator at this time.
When you are comfortable with these instructions...
click the green button below that says I'm done with the instructions. You will be asked some questions that test your understanding of these instructions. Once you answer them correctly, your task will begin.
I'm done with the instructions
Feel free to use the << Previous page and Next page >> buttons to review the instructions while you are answering the questions.
<< Previous page
Next page >>

2-9. Appendix B – Instructions for main round

This appendix shows the instructions for the main round. These instructions were displayed one page at a time using a viewer application written in the C# programming language. The viewer displayed the instructions in a panel on the left, and a sample screen, if available for that page of the instructions, was shown in a separate panel on the right. The panel at right was also used to administer comprehension and manipulation check questions for the main round. These questions were presented to participants after they clicked the I'm done with the instructions button on the final page of the instructions.

Different main round instructions were required for each experimental condition. To implement this, I created master document with embedded tags, such as the **TT** and **PPPLUS** tags shown in pages 2 and 3 of the instructions below. At runtime, the viewer application parsed the master document and replaced each tag with the appropriate value for the experimental condition for that participant. The table below shows the value of each tag for each experimental condition:

Tag	Frequent feedback Low uncertainty	Frequent feedback High uncertainty	Infrequent feedback Low uncertainty	Infrequent feedback High uncertainty
TT	The tag **TT** was the participant's point target. The point target was computed as follows. For each participant, the target was set to the number of tables counted by that participant in the second half of the practice round, multiplied by 2. This number was multiplied by 1.175 (increased by 17.5%). This represented the number of tables the participant could count in a 10-minute period if they work at a rate 17.5% higher than in the second half of the practice round. This number was multiplied by 10 points per table to obtain a point target. It was then rounded to the nearest 10. The minimum allowed value of the target was 410, and the maximum was 550. Thus, *the target was customized for each participant, and did not depend on experimental condition*.			
PPMINUS	8	2	8	2
PPPLUS	12	18	12	18
SUM	42	48	42	48
MIN	1	1	4	4
MIN_STR	“”	“”	s	s

Instructions - page 1 of 9

Now that you have some experience with the task, there are some changes. But first, let's review what has not changed.

What has not changed

- Your task remains the same – counting zeros.
- You will still have three tries to count each table, you will be assessed a penalty of 10 points for three wrong answers.

What has changed

You will be paid differently for the remainder of the session. Click the Next page >> button for the details.

<< Previous page | Next page >>

Instructions - page 2 of 9

Three periods

- **You will work for three 10-minute periods.**
- Each 10-minute period is independent of the others.
  - ❑ You will begin each period with zero points.

Points targets

- **If you earn {TT} or more points during a period, your pay will increase by $5.**
- Examples:
  - ❑ If you earn {TT} or more points in *only one* of the three periods, your pay will increase by $5.
  - ❑ If you earn {TT} or more points in *two* of the three periods, your pay will increase by $10.
  - ❑ If you earn {TT} or more points in *all three* of the periods, your pay will increase by $15.
- There is no bonus for exceeding {TT} points during a period.
- Points earned during one period do not carry forward to the next period.

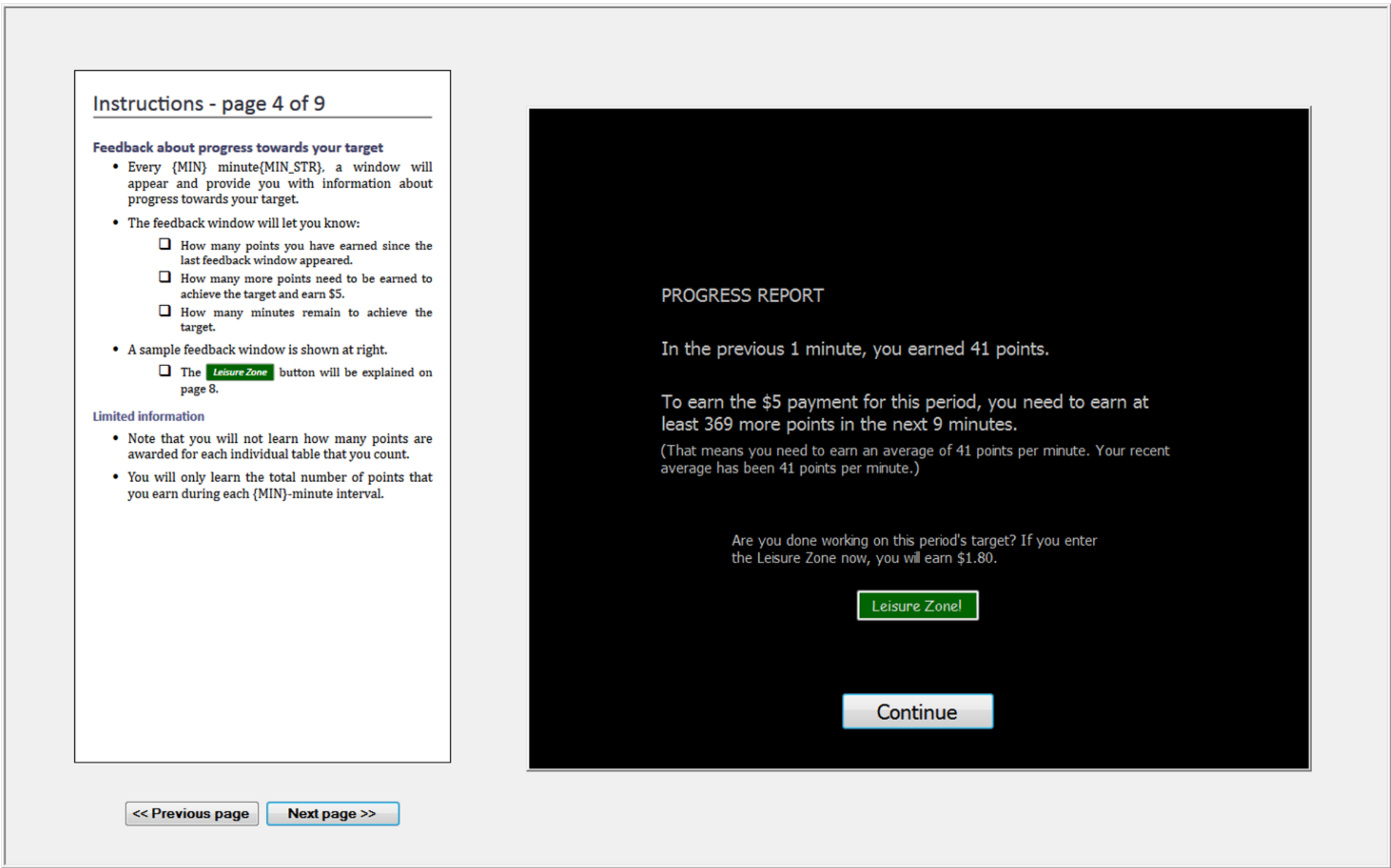
<< Previous page | Next page >>

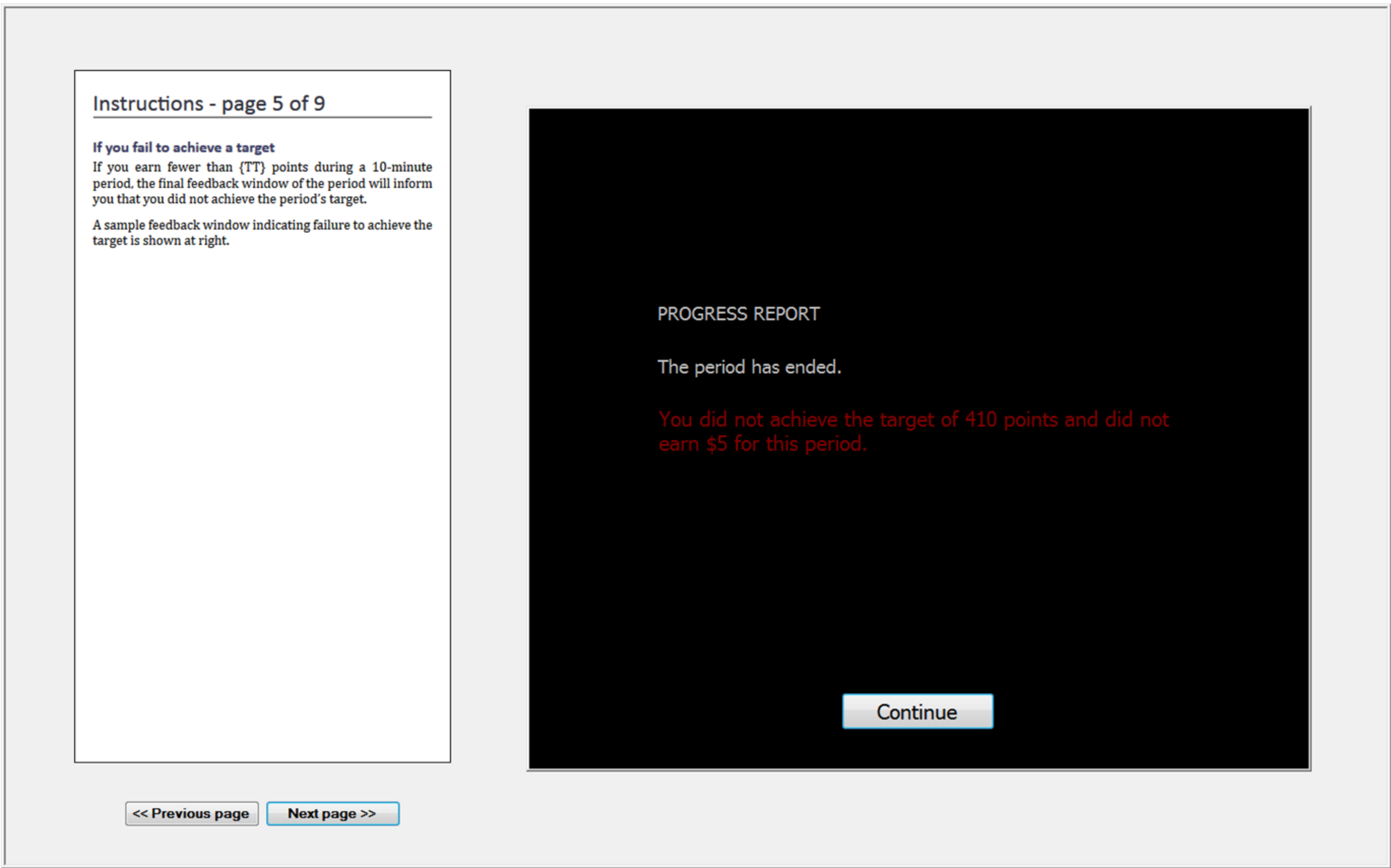
Instructions - page 3 of 9

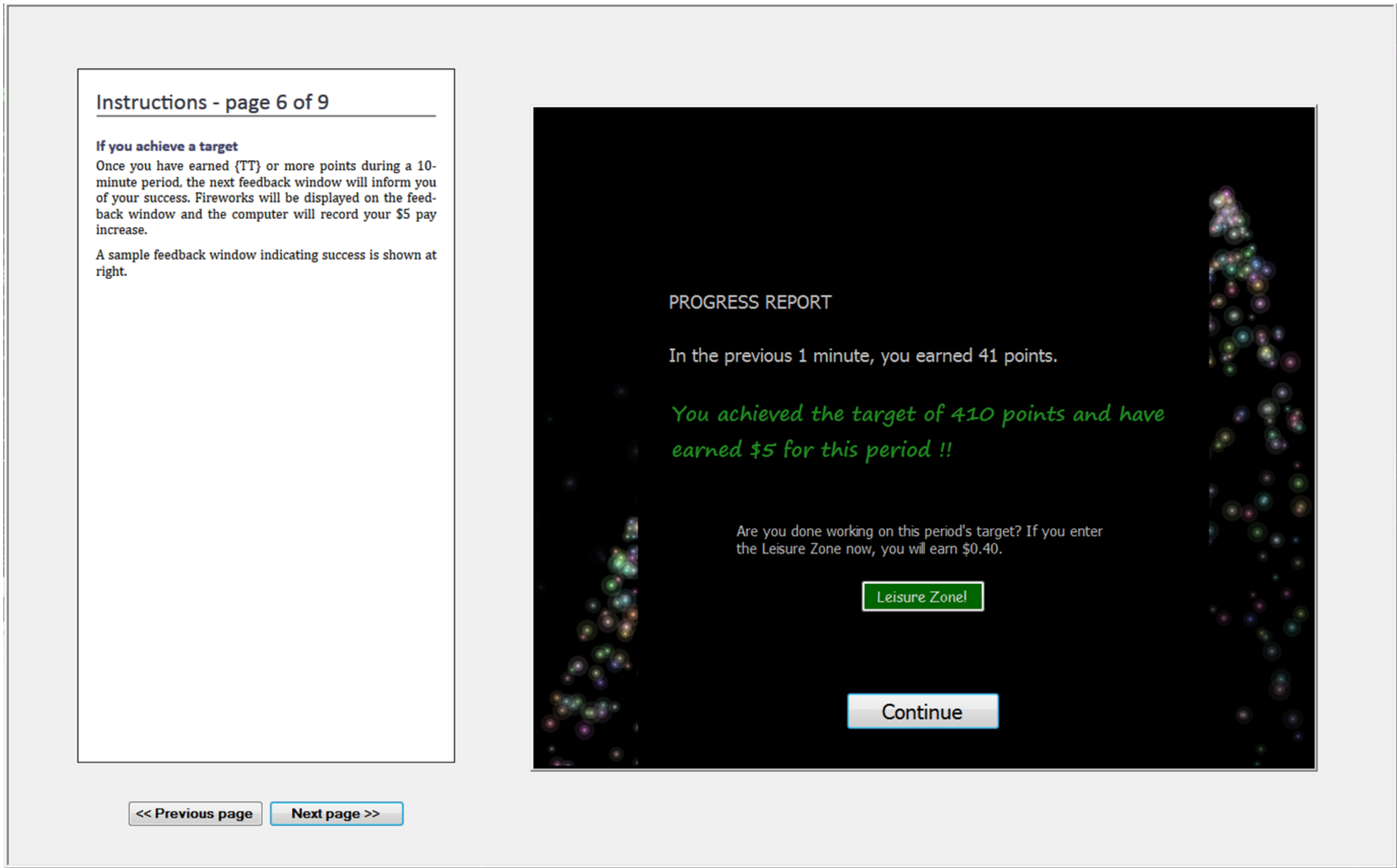
Points per table

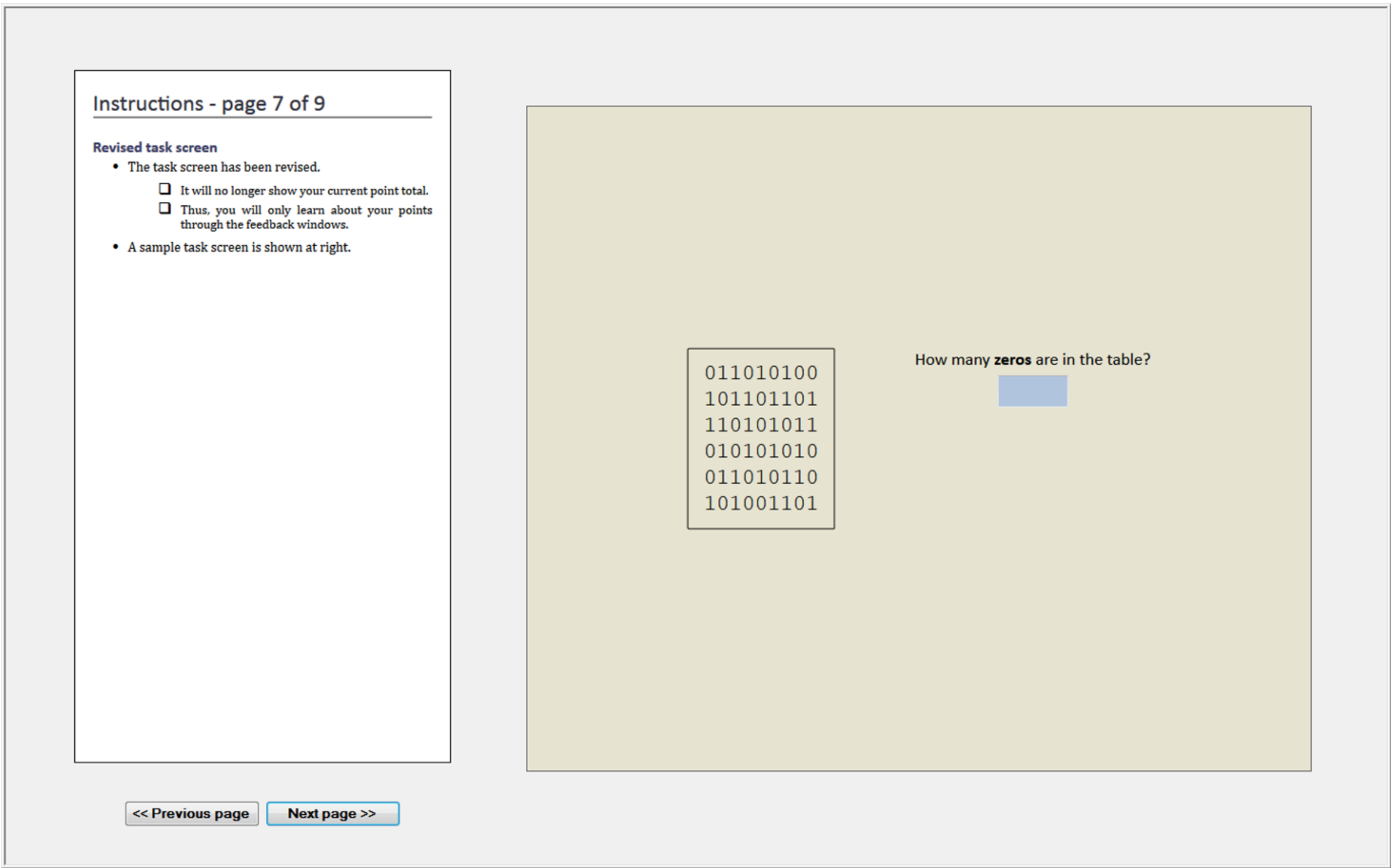
- **For each table that you count correctly, you will be awarded {PPMINUS}, 10, or {PPPLUS} points.**
- The point award will be chosen independently *for each table*.
- The average point award over all three periods will be approximately 10 points.
- Example:
  - ❑ You count four tables.
  - ❑ You are awarded {PPPLUS} points for the first table, {PPPLUS} points for the second table, {PPMINUS} points for the third table, and 10 points for the fourth table.
  - ❑ Thus, for counting these four tables, you earn a total of {PPPLUS} + {PPPLUS} + {PPMINUS} + 10 = {SUM} points.

<< Previous page | Next page >>


Instructions - page 4 of 9
Feedback about progress towards your target
Every {MIN} minute{MIN_STR}, a window will appear and provide you with information about progress towards your target.
The feedback window will let you know:
How many points you have earned since the last feedback window appeared.
How many more points need to be earned to achieve the target and earn $5.
How many minutes remain to achieve the target.
A sample feedback window is shown at right.
The Leisure Zone button will be explained on page 8.
Limited information
Note that you will not learn how many points are awarded for each individual table that you count.
You will only learn the total number of points that you earn during each {MIN}-minute interval.
PROGRESS REPORT
In the previous 1 minute, you earned 41 points.
To earn the $5 payment for this period, you need to earn at least 369 more points in the next 9 minutes.
(That means you need to earn an average of 41 points per minute. Your recent average has been 41 points per minute.)
Are you done working on this period's target? If you enter the Leisure Zone now, you will earn $1.80.
Leisure Zone!
Continue
<< Previous page
Next page >>


Instructions - page 5 of 9
If you fail to achieve a target
If you earn fewer than {TT} points during a 10-minute period, the final feedback window of the period will inform you that you did not achieve the period's target.
A sample feedback window indicating failure to achieve the target is shown at right.
<< Previous page
Next page >>
PROGRESS REPORT
The period has ended.
You did not achieve the target of 410 points and did not earn $5 for this period.
Continue


Instructions - page 6 of 9
If you achieve a target
Once you have earned {TT} or more points during a 10-minute period, the next feedback window will inform you of your success. Fireworks will be displayed on the feed-back window and the computer will record your $5 pay increase.
A sample feedback window indicating success is shown at right.
<< Previous page
Next page >>
PROGRESS REPORT
In the previous 1 minute, you earned 41 points.
You achieved the target of 410 points and have earned $5 for this period !!
Are you done working on this period's target? If you enter the Leisure Zone now, you will earn $0.40.
Leisure Zone!
Continue


Instructions - page 7 of 9
Revised task screen
The task screen has been revised.
It will no longer show your current point total.
Thus, you will only learn about your points through the feedback windows.
A sample task screen is shown at right.
011010100
101101101
110101011
010101010
011010110
101001101
How many zeros are in the table?
<< Previous page
Next page >>

Instructions - page 8 of 9

The *Leisure Zone*

- Whenever a feedback window is open, you can choose whether you want to continue counting tables, or whether you want to browse the web for the remainder of the 10-minute period.
- If you decide to browse the web, press the *Leisure Zone* button on the feedback screen.
- You may visit any web site you like.

Pay for time spent in the *Leisure Zone*

- **You will be paid $0.20 for every minute that you spend in the *Leisure Zone*.**
- If you have already achieved the target of {TT} points when you enter the *Leisure Zone*, your pay for time spent in the *Leisure Zone* will be in addition to the $5 for target achievement.

Restriction

- Once you enter the *Leisure Zone*, you cannot leave it until the current 10-minute period is over.
  - ❑ Thus, you will not be able to count any more tables in the current period after you enter the *Leisure Zone*.
  - ❑ If you enter the *Leisure Zone* before achieving the period's target, you will not be able to achieve the target.
- Each period is independent. Entering the *Leisure Zone* in one period will have no effect on future periods.

<< Previous page | Next page >>

Instructions - page 9 of 9

If you have questions...
please ask the lab administrator at this time.

When you are comfortable with these instructions...
click the green button below that says *I'm done with the instructions*. You will be asked some questions that test your understanding of these instructions. Once you answer them correctly, your task will begin.

I'm done with the instructions

Feel free to use the << Previous page and Next page >> buttons to review the instructions while you are answering the questions.

<< Previous page   Next page >>