Is "Real" Effort More Real?*

E. Glenn Dutcher[†] Timothy C. Salmon[‡] Krista J. Saral[§]

September 2020

Abstract

A growing number of studies use "real" effort designs for laboratory experiments where subjects engage in an actual task as a form of effort provision. The commonly argued reason for using real effort is that it is more generalizable and field relevant than using stylized effort where subjects simply choose some level of costly effort. We find little support for these claims. Moreover, we revisit past real effort designs and note that a common thread is that they lack any cost of effort. To make this point clear, we first provide a theoretical characterization of the nature of effort costs to provide a better understanding of how to model them in the lab. We then conduct two experiments with the aim of testing whether there are differences in behavior between real and stylized designs. In both tests we find that there are no differences though we also discuss situations in which that might not be the case. We also design and test a module for implementing proper effort costs in a real effort experiment which could be added on to almost any real effort design which should help overcome some of the shortcomings we demonstrate with standard real effort experiments.

JEL Codes: C91, H41 **Key Words:** Real Effort, Stylized Effort, Abstract Effort, Economics Experiments, Public Goods, Slider Task, Incentives, Coordination Games

*The authors would like to thank Jimmy Walker and Mark Isaac for providing the instructions from their early papers on public goods games, and Lise Vesterlund for providing the instruction scripts, software programs, and data we used in the design of one of our treatments. We would also like to thank participants at the ESA World Meetings in Vancouver, ESA North American meetings in Dallas and the Texas Experimental Conference at Rice for many useful commends and suggestions.

[†]Ohio University, Department of Economics, Bentley Annex 3rd Floor, Athens, Ohio 45701. dutcherg@ohio.edu

[‡]Southern Methodist University, Department of Economics, 3300 Dyer Street, Suite 301 Umphrey Lee Center, Dallas, TX 75275-0496. tsalmon@smu.edu, Phone: 214-768-3547, Fax: 214-768-1821.

[§]University of North Carolina at Charlotte Department of Economics, 9201 University City Blvd. Charlotte, NC 28223, USA; Webster University Geneva George Herbert Walker School of Business and Technology, Geneva, Switzerland; GATE, Ecully, France. E-mail: ksaral@uncc.edu

1. Introduction

There are many types of economic experiments that involve having subjects put forth some form of "effort." While effort has been proxied in different ways, the approaches are generally classified into two different categories: stylized (or chosen effort) and real effort. Stylized effort involves subjects choosing a number from a prespecified range to represent their effort level with higher numbers being more financially costly than lower ones. Real effort designs include a wide range of options in which subjects perform some task as their form of effort which could be solving mathematics problems, solving mazes or other puzzles, stuffing envelopes, or even in one case shelling walnuts (Fahr and Irlenbusch (2000)). The choice of whether to use real or stylized effort is potentially important as it could have substantial implications for the outcome of the experiment. In this paper we examine the foundations of modeling effort provision in experiments to better understand the important issues in the choice of whether and how to implement stylized or real effort in experiments.

Both approaches to representing effort in the laboratory are attempts to replicate important elements of effort decisions in field settings. Stylized effort is arguably better suited to capture theoretically derived trade-offs present in field effort choices. Any trade-off between effort and other options that one believes to be important to a field setting can be specified in a theoretical structure and implemented in a stylized design. The theoretical predictions can then be tested directly in the laboratory by simply giving subjects the same choice faced in the model, which permits careful testing of whether or not subjects' choices comply with theoretical predictions. There are of course potential limitations to stylized designs as it is possible that they could omit elements of effort provision that are important in the field but not (explicitly) contained in the corresponding theoretical model. It is also possible that stylized designs leave out some contextual issues important in field settings.

The primary argument in favor of the real effort approach is that it involves subjects engaging in actual physical or mental effort to complete a task, just as one would do in a field setting. Real effort designs are often assumed to be better at capturing trade-offs present in field situations due to the fact that actual effort is involved. There are innumerable claims found in the literature about how eliciting effort using these tasks leads to results that are inherently more generalizable to the field than stylized effort designs. The most direct statement we found to this effect is in Gill and Prowse (2011): "The main advantage of using a real effort task over a monetary cost function is the greater external validity of the experiment: exerting actual effort makes the environment more realistic and less sterile, increasing the likelihood that the motivations that drive behavior outside the laboratory carry over to the laboratory."(p.1) This point is reinforced in a recent survey regarding real effort experiments, Charness, Gneezy, and Henderson (2018), in which the authors point out that "The advantage of the real-effort method is that it is closer to the psychology of working. For example, the cost of effort might vary over time: solving mazes might be fun initially, but might gradually become less motivating." (p.75) The basic implication from these types of claims is that by its nature, a real effort task has a cost function that is a better representation of the effort cost in external situations. While such claims are intuitively appealing and quite common in the literature, we have found no evidence of any theoretical or empirical support for them.

In order to empirically examine claims regarding the superiority of real effort designs in replicating a field environment, two things must be demonstrated. First, given comparable real and stylized effort decisions, people make different choices. Second, the choices made using the real effort version are a better match with field behavior. While we are not aware of studies addressing the second, its relevance is conditional on the first being true. We have found several studies trying to address the first issue but they typically do not hold effort costs constant between the two designs. In the stylized designs, there is an explicit monetary cost of choosing higher effort levels. In real effort designs, effort costs are discussed but rarely does a real effort study explain what is being given up to engage in the effort in the experiment or to put it differently, why effort is costly. Even rarer is a study that finds a way to actually keep effort costs constant between the two designs. To compare real and stylized designs, we must first come to a more coherent understanding of what effort costs actually are in real effort studies so that it is possible to construct real and stylized effort designs that have the same effort cost.

We present a theoretical characterization of a standard labor supply model to demonstrate what an effort cost is and where it comes from; i.e., an effort cost is an opportunity cost. The reason effort is costly to a worker is that to engage in effort provision on the main task, the worker must give something up. What they give up is the opportunity of spending that same amount of time, energy, mental focus, etc. . . on a different activity. If a worker has an alternate activity to engage in which yields positive utility, effort is costly. If they do not, then they lose nothing by expending the effort on the work activity, meaning the cost of effort is zero.

While this realization comes from any basic model of labor supply, it is an important insight that is often left out of real effort designs. Typically in these experiments, subjects are given the opportunity to engage in some task with their output on that task labeled as "effort" and they are given no alternative activity to engage in. This means their only actual alternative is staring idly at the computer screen. There are existing studies showing that subjects in experiments would rather subject themselves to very painful electric shocks than sit idly for a few minutes (Wilson, Reinhard, Westgate, Gilbert, Ellerbeck, Hahn, Brown, and Shaked (2014)), and many real effort experiments have unintentionally replicated this finding. Araujo, Carbone, Conell-Price, Dunietz, Jaroszewicz, Landsman, Lamé, Vesterlund, Wang, and Wilson (2016) replicate it intentionally as they examine a commonly used real effort mechanism and demonstrate that increasing the piece rate 1600% from $0.005 per unit of output to $0.08 has no impact on output. The authors conclude that there is a problem with the specific real effort task they implement but we will argue that the task itself may not actually be the problem - the same result of a lack of a wage effect appears in many other real effort papers using different tasks such as a spelling test in Esarey, Salmon, and Barrilleaux (2012) and a standard encoding task in Ku and Salmon (2012). The lack of a wage effect in these cases is easily explainable by the fact that there is no cost of supplying effort in these designs. Thus, the lowest wage offered compensates the subjects adequately for spending the entire time allotment on the task, and higher wages can induce no greater effort, or time spent on the task.

Erkal, Gangadharan, and Koh (2018) present a careful documentation of the importance of an outside option as the authors conduct a variety of experiments that have the effect of changing the value of the outside option to examine the impact on behavior. For example, they conduct sessions in which subjects must stay in the laboratory for a certain period of time while in others, they can leave at a time of their choosing. In the latter, a subject has a cost of effort because by staying they are giving up the option of engaging in some other activity. In the former, there is no cost. Not surprisingly, measured effort is lower in the case with an actual effort cost. While these authors identify this regularity, they do not fully investigate it or clearly identify the theoretical mechanism behind it. It is important to do so since by understanding why such results keep appearing in the literature we will be able to better understand how to design future experiments and to interpret results from experiments.

In some real effort designs, the lack of any actual effort cost may not have any impact on the eventual results. In others, the lack of an effort cost can cause serious inference mistakes. One example of this inference problem is given by Engel (2010). In that paper the author notes that prior studies had found individuals were willing to work above thresholds required for payment but these were also cases where additional effort was not costly. These prior papers had concluded that the effort above the required level was due to some intrinsic motivation or moral commitment to work. Engel (2010) conducts new experiments where subjects could switch to an alternate activity once the threshold for payment was met and found that they generally did so. Thus it seems reasonable to conclude that the prior explanations for the excess effort were reached in error due to the design of the experiments. Corgnet, Hernán-González, and Schniter (2015) examine a team production setting using real effort with and without an outside option and they also find the results differ substantially depending on whether the outside option is included in the experiment.

After our theoretical investigation of the foundations of effort costs we conduct three different experiments to examine different issues that arise in examining real and stylized effort designs. Our first experiment examines some prior claims regarding a domain in which real and stylized effort led to different behavioral outcomes. In our experiment we re-examine this claim but hold effort costs constant between the two designs and find that behavior does not differ between the two cases suggesting that a better explanation of the prior results is that they were due to differences in effort cost between the real and stylized designs, not due to the modality of effort provision itself. In our second experiment, we propose and then test a mechanism for adding effort costs into any real effort experiment. The mechanism we propose could be easily added into just about any real effort experiment to allow an experimenter to control the costs of effort. We investigate the capability of this mechanism by reconducting the tests done in Araujo et al. regarding whether or not a labor supply response can be observed when using the slider task in a real effort experiment. Our results demonstrate that when actual effort costs are incorporated into the slider task, effort/output does increase with piece rate wage as expected. In our final experiment, we directly test the proposition that one should expect different behavior in real and stylized designs. This last experiment is informed by the lessons of the prior two experiments. We design an environment in which the only thing that differs between the real and stylized designs is the effort modality, and we do so in a context where it is reasonable that real effort could trigger something different in the preferences of the subjects, leading to different behavior. We again find no differences in behavior, though we do not intend to extrapolate this result to suggest that such differences cannot exist. We return to this point in our conclusion where we provide a clearer statement regarding how we think these results should be interpreted.

2. Cost of Effort

2.1. Theory

To understand the origin of what an effort cost is, we start from a standard version of the neoclassical model of labor supply. This model involves an individual deciding how to divide a fixed time budget between multiple activities. In the classic version, there are only two activities, work and leisure. One could expand the model to include any number of activities for the individual to divide time between but we will use the two-activity version as that is enough to demonstrate the key trade-offs involved with time allocation. We assume that an individual receives utility from consumption and from time spent engaging in leisure. For simplicity we assume the two sources of utility are additively separable. Let $W(y)$ be the utility an individual derives from consumption which is a direct function of the earnings, y, they receive from time spent working. The earnings an individual receives, y, will be assumed to be a function of the time allocated to the work activity, x, where x is what one usually means when they refer to the effort of a worker. Let ℓ represent the amount of time spent on leisure leading to a utility of $L(\ell)$ where leisure does not imply inactivity, merely time spent not working. As a terminology note, we will use the term "effort" to refer to time spent on the work activity even though many people exert a great deal of effort on their leisure activities (e.g. marathon running, mountain climbing, etc. . .). Even though this is the conventional use of the term effort, we emphasize this convention to prevent confusion. While individuals may be active in either their work or leisure pursuits, the term effort will be reserved only for activity directed towards the work activity even if an activity specified to take the place of the leisure activity is one that would also require some degree of active engagement.

Based on these assumptions, a worker has an overall utility function $U(x, \ell) = W(y(x))+$ $L(\ell)$. It is standard to assume that utility is increasing in both arguments or $W_y > 0$ and $L_\ell > 0$. As time is a scarce resource, there is a fixed amount, T, the worker can choose to divide between these two utility earning activities which leads to the standard labor supply problem

$$\max_{x,\ell} W(y(x)) + L(\ell) \tag{2.1}$$
$$s.t.\ x + \ell \leq T$$
$$x \geq 0, \ell \geq 0$$

Given the non-satiation assumptions, i.e. $W_y > 0$ and $L_\ell > 0$, and the typical assumption that earnings increase with time spent working, $y_x \geq 0$, the inequality constraint is always binding and so the optimality condition for this problem is simply

$$W'y' = L' \tag{2.2}$$

or the marginal benefit of labor due to the increased consumption possible from more time spent working, $W'y'$, is equal to the marginal cost of effort, L'. The clear interpretation is that allocating another unit of time to labor is costly because this requires taking that unit of time away from the leisure activity. The cost of one more unit of labor is therefore equal to the utility decrease from giving up spending that amount of time on leisure.

Many experimental studies that involve effort provision do not explicitly use the standard labor supply model as their foundation, rather they use some variant of a standard principal-agent model. The specification is simply that the agent earns utility from income which is increasing in effort, but the agent also experiences disutility from expending the effort. A simple linearized version of this model would specify the following choice problem

$$\max_{e} y(e) - c(e) \tag{2.3}$$

where $y(\cdot)$ is some function indicating how effort yields income and $c(\cdot)$ measures the cost of effort, normalized into monetary value. The origin of $c(\cdot)$ is generally not specifically addressed in papers using this model as it is simply some function assumed to have positive first and second derivatives. From a theoretical perspective, that's a useful specification. However, once we try to determine the empirical and experimental representation of $c(\cdot)$, we have to think more deeply about where it comes from. Why is effort costly? The labor supply model provides the answer and it is because one must give up something to engage in that effort.[1]

To understand what is given up by the agent in the principle-agent model, we need to connect it back to the labor supply model. Effort or e in the principal-agent model is typically not denominated in time spent on the effort task but rather it is an abstract notion of effort. In real effort experiments it is usually measured in terms of tasks completed. Both of these are useful simplifications but they are simplifications. In theoretical models, one may prefer to abstract away from specific time frames and production functions. In experiments, measuring effort in completed tasks allows for a convenient definition of payment and cost functions. However, this approach is really a shortcut as a subject in

[1]For example Mas-Colell, Whinston, and Green (1995), chapter 13 p. 438, identifies this term as foregone earnings from home production which is simply another way to note that it is an opportunity const.

an experiment cannot actually choose an arbitrary number of math problems to complete in 4 minutes. What they can choose is how much time to spend trying to solve math problems. This time spent is converted to output based on some underlying production function that depends on their capability for solving math problems. Thus at the problem's core, time, not the number of completed problems, is still the choice variable as it is the scarce resource being divided. The reason completing more math problems is costly is because time spent doing so cannot be allocated to the next best option, whatever that may be.

This description of effort costs does ignore the possibility that effort may involve dimensions in addition to time. One way of conceptualizing other cost dimensions is to refer to them collectively as the intensity dimension of an individual's work effort. An individual choosing to work with greater intensity during a 3 minute time period could solve more math problems than if they were to choose to work with lower intensity for the same amount of time. Outside the laboratory, this intensity choice is likely to be very important as a manual laborer who works with great intensity for 4 hours may exhaust themselves to the point that they diminish how much enjoyment they receive from their leisure time, while a worker who chooses to work with less intensity may still be rested and able to enjoy their leisure time.[2] Inside the laboratory, subjects can certainly vary the intensity of how much they focus on a task and we could model this with an additional choice constraint where the subject is also allocating some other scarce resource such as mental focus. It is not clear that this is necessary as it strains credulity to think that working a little harder on a math task in an experiment diminishes the enjoyment of staring at a screen - the only other alternative available in many studies - due to fatigue or that the extra focus inside of the experiment leads to substantial fatigue post-experiment. Subjects may vary their work intensity due to various contextual elements in the experiment (e.g. mood) and that can be an important aspect of an experiment design and research question, but it is unlikely that this increase in intensity in a short production period is due to subjects wishing to save the energy costs of aligning a slider or solving a math problem. In fact, if these intensity costs were substantial the results previously cited showing a lack of a wage effect would not have been found. Those results make it clear that whatever these intensity costs are, they are less than $0.005 per slider, and so it seems reasonable to round them down to zero for typical real effort experiments.

[2] There is a long literature on "ego depletion" which would be a very noted example of this idea as it supposes that an individual has a limited budget of mental resources and when those are used up, it causes a person to behave with low self-control. Those conclusions are currently in substantial dispute, see Hagger et al. (2016), but we note this as a well-known example of a model in which the scarce resource is a version of mental energy.

Given the intuitive link between the two models, it is straightforward to show the connection between the standard labor supply model and the principal agent model. To do so, let us redefine the choice problem in the principal agent model to be one of time division. Instead of choosing output, e, we will assume the individual chooses time to devote to the productive task, x, and $e(\cdot)$ becomes a production function which translates time spent on the task to output. With this structure, we can model our agent as choosing how much time, x, to devote towards the labor task out of a total time budget of T. If the agent spends their entire time budget on production, they produce $e(T)$ units of output. If they spend some time $x < T$ on production then the difference $T - x$ indicates the time spent on the leisure activity. Assuming our time constraint holds as an equality, we can define $L(T - x)$. Further, we can define $L(T) = H$ as a constant representing the maximum leisure utility possible from spending no time on the productive task. We can then rewrite $L(T - x) = H - f(x)$ where $f(\cdot)$ measures the decline in leisure utility due to allocating time x to production. Therefore $f(x)$ measures the cost of spending time x on the productive task. With these definitions, we can rewrite the labor supply model, Equation 2.1, as

$$\begin{gathered} \max_x W(y(e(x))) + H - f(x) \\ s.t.\ x \leq T \\ x \geq 0 \end{gathered} \tag{2.4}$$

It should be clear that equation 2.3 is a simplified version of equation 2.4. First, since H is a constant, it can be dropped out without affecting the first order condition or comparative statics so leaving it out of 2.3 has little impact. Second, it is clear that $c(e(x)) = f(x)$ as both measure the decrease in utility from spending time on the effort task. Further, the utility of earnings from labor is simplified to assume that the utility of the earnings is equal to the amount earned or $W(y(e(x))) = y(e(x))$. And of course the final simplification involves abstracting away from the specific resource being divided and simply calling the choice "effort" rather than stipulating that effort is measured in time, energy, mental focus, or something else. The main point is that the principal-agent representation is just a simplification of the standard labor supply model.

Understanding the connection between these models is important as it provides a key insight on the nature of effort costs. First, as stated earlier, all effort costs are opportunity costs. They represent decreases in utility from not spending the relevant resources on the next best available activity. A second key insight is that the cost of effort is unrelated to the nature of the effort task itself. The cost of effort derives from the utility an individual would receive from engaging in their alternative activity. The effort task affects this only through the resources, say time, not spent on it. So if the next best alternative is time spent watching a movie, whether the hour spent on the effort task is spent sorting papers or shelling walnuts, the effort cost is the same because the individual has given up the utility they could have received from spending that hour watching a movie. The specific nature of the effort task can affect the ability of the individual to divide time between the tasks, but other than this, there is no impact of the nature of the productive task on the effort cost. This is contrary to the standard view in the literature which is that the cost of effort derives from the task itself – a point which is summarized in Charness, Gneezy, and Henderson (2018), "Control over the cost-of-effort function, seen as one of the major advantages of the chosen-effort paradigm, has been addressed primarily through qualitative means, for example by juxtaposing results from 'easy' and 'hard' real-effort tasks." Understanding that effort costs are derived from the outside option is also an important insight in regard to the claimed field relevance of the real effort tasks. In the field, the reason a worker faces a cost of effort is because if they are not engaged in the productive activity for their job they could be engaging in a broad range of activities including working for a different employer, watching a TV show or movie, chatting with friends, online shopping, riding a dirt bike or any other pursuit. Unless these outside options are included in a real effort experiment, it is quite difficult to claim that the cost of effort reflects effort costs outside the laboratory. Of course including these types of outside options in a laboratory experiment is difficult to do in a practical manner which means other approaches must be taken to incorporate effort costs into the laboratory environment. An additional implication is that any approach to inducing effort costs into a real effort experiment will still involve implementing an abstract notion of an effort cost intended to match these effort costs from the field.

We do not mean to imply that the nature of a real effort task cannot affect behavior, but rather our goal is to focus on identifying what aspects of a task can affect behavior and how best to represent them both in a model and in a corresponding experiment design. For example, some individuals may find some tasks more enjoyable than others which will make them willing to pay more of a cost to engage in the task than others. Similarly, the production function is unlikely to be the same across people as some may be faster or slower at a task than others. Such elements are commonly discussed and when these elements are included in a model they are often included as modifications to the cost function rather than as modifications to the utility function for utility earned from completing the task or to the production function. One can certainly write down a model which yields identical results between putting the ability differential in a production function or a cost function and often it may be theoretically more convenient or tractable to write the theoretical model in this manner. While convenient and often unimportant for examining certain theoretical comparative static responses, it is a misspecification that can and has led to researchers misunderstanding the nature of effort costs and constructing experiment designs inappropriately. The problem is that by placing these elements in the cost function, one is confounding lower ability to solve math problems with a greater appreciation of leisure time, or disliking mathematics problems with increased utility from leisure time. There is no clear reason for there to be a connection between how much one appreciates leisure time and the nature of the effort task or one's ability at it. While this misspecification may not have substantial complications for purely theoretical applications, it can cause problems when experimenters construct environments which match those theoretical constructions. Understanding the proper place for these elements in the theoretical specification allows a researcher to design an experiment that better maps into the field relevant situations.

The key insights from this section are important for understanding how to model effort in an experiment. The most important point is that all effort costs are opportunity costs and if effort in an experiment is to be costly, a subject must give up something of quantifiable value in order to engage in it. In the next sections, we examine the importance of this insight in interpreting prior comparisons between real and stylized effort; we also demonstrate a means to include effort costs as a simple add-on to any real effort design. Last, we engage in a careful test of the possibility that real and stylized effort may yield different behavior even after carefully controlling for effort costs and other differences between the designs.

3. Experiment 1 Design: Verifying the Importance of Effort Costs

In this first experiment we aim to understand why some previous investigations of real and stylized effort have shown differences in behavior. The insight from our theoretical investigation above suggests that one reason for these differences could well be that effort costs between the two designs were not the same. To investigate this point, we will examine the experiments in Bortolotti, Devetag, and Ortmann (2009) which suggest that individuals are more willing/able to coordinate when the task is a real effort task rather than stylized effort. We chose this set of experiments for our investigation due to the fact that its methods are fairly standard in literature and that the substantive claims made are important to understand. In this paper, the results from a standard stylized weak link coordination game are compared to a real effort study in which individuals counted coins and were paid by the lowest error rate among the members of their group. While the stylized coordination game has costs built directly into it to ensure that effort above the minimum is costly, it is not clear that there were significant effort costs to the coin counting exercise because this is the only task subjects could engage in.[3] This is an important difference between the real and stylized designs as this difference in effort costs could explain the difference in results rather than the results being attributable to a fundamental difference between real and stylized effort.

Similar differences in effort costs using stylized designs were already examined in Van Huyck, Battalio, and Beil (1990) (VHBB). In that study the authors conduct two versions of coordination games, one in which it was costly to contribute above the minimum contribution in the group and another in which it was costless. In the version where it was costly to contribute above the minimum of others, coordination failed with groups ending up at the minimum contribution level. On the other hand, when contributing above the minimum was not costly, 96% of all subjects chose to coordinate on the highest choice. This is a stark difference and suggestive of the differences found in Bortolotti, Devetag, and Ortmann (2009) between their real and stylized effort designs. Given that there were two element changes between the treatments in Bortolotti, Devetag, and Ortmann (2009), the effort modality and the effort cost, we want to examine which drove the difference in the ability to coordinate. To that end we conduct an experiment with a real effort coordination game in which we vary whether or not there are effort costs to determine if our results match with what VHBB found in their stylized experiments. If our results in a real effort setting match those in VHBB, it will support the claim that the differences in behavior found in Bortolotti, Devetag, and Ortmann (2009) are due to the effort cost difference and not the difference in effort modality. On the other hand, if we find that subjects are able to coordinate well in the real effort coordination regardless of effort costs then it provides support for the claim that the effort modality drives the original results.

In our real effort version of a coordination game, subjects participate in four person groups and have the opportunity to complete instances of a task for earnings in one of two between-subjects treatments: Costless Effort and Costly Effort. The task they engage in is a counting task where subjects are asked to count the number of 0's in a string of 0's and

[3]The authors, to their credit, included an option on the real effort experiment which would allow subjects to buy extra time to complete their task which could be seen as a clear effort cost. After an initial learning phase, subjects generally seemed to have no need of the extra time meaning it was rarely an actual cost. Even with this cost in place though, it isn't clear that this cost function would have been calibrated to represent the same cost as in the stylized design.

1's. Subjects can apply correctly completed tasks toward a team account or (possibly) an individual account to generate earnings. Linking back to the theoretical discussion, tasks directed to the team are usually thought of as "effort" towards work output and the effort directed towards the individual account can therefore be viewed as leisure despite the fact that the subject is still engaging in active work in their leisure activity.[4] The team account's earnings follow the same structure as in the VHBB weak link game; they receive piece rate earnings based upon the lowest number of tasks directed to the group account by any member of their group. When effort is costly, subjects also earn a piece rate based on how many tasks they complete and direct towards their individual account. When effort is costless the individual account is eliminated from consideration, meaning the subject can only contribute to the team account. This is equivalent to receiving no compensation for tasks directed to the individual account and thus there is no opportunity cost for effort directed towards the group account. We eliminate the option rather than set the compensation to 0 as this better mimics how standard real effort experiments are conducted. In all treatments we also include a fixed payment per period.

The costly treatment is mathematically equivalent to VHBB. In VHBB, the payoff function is

$$\pi_i^1(e_i, e_{-i}) = f + b * \min(e_i, e_{-i}) - c * e_i \tag{3.1}$$

where f is a fixed payoff, b is the per unit payment for the minimum of the effort choices among group members, and c is the cost of the individual effort choice. In our version, we keep the same incentives for the group task in that subjects receive earnings of $b * \min(e_i, e_{-i})$ or b times the minimum number of tasks completed by a member of their team. Each task directed towards the team account, however, implies they must give up earnings from the individual account. If we let l_i be the number of tasks directed toward the individual account and pay c per task, then total earnings are $b * \min(e_i, e_{-i}) + c * l_i$. Given the limited time to complete the tasks, there is a trade-off between e_i and l_i meaning that $e_i + l_i = T$ or $l_i = T - e_i$. Substituting $c * l_i = c * T - c * e_i$ gives earnings as $c * T + b * \min(e_i, e_{-i}) - c * e_i$. Because $c * T$ is a constant it can be seen as a part of the fixed payoff, f , in the original VHBB payoff function giving us a slightly transformed

[4] Methodologically, it is important that the tasks to be completed for the individual and group accounts are the same. It makes it easier to measure the effort costs this way as it should take the subjects the same time to complete an instance of this counting task regardless of which account they direct it to. Thus the only difference is the piece rate earnings. If the tasks were different, then we would have to account for the difference in time to complete the two tasks in measuring how much a subject gives up on earnings from one task to complete one unit of the other.

		Minimum of the Team's Contribution							
		7	6	5	4	3	2	1	0
Own	7	200	180	160	140	120	100	80	60
Contribution	6		190	170	150	130	110	90	70
	5			180	160	140	120	100	80
	4				170	150	130	110	90
	3					160	140	120	100
	2						150	130	110
	1							140	120
	0								130

Table 3.1: Coordination game with effort cost.

version of their original function written as

$$\pi_i^2(e_i, e_{-i}) = (f' + c * T) + b * \min(e_i, e_{-i}) - c * e_i \tag{3.2}$$

So long as $f = (f' + c * T)$ then these payoff functions are identical. What this transformation highlights is that the cost of one more unit of production directed towards the group is a decrease in individual earnings by c, which is the same costs faced in VHBB and shows, again, the link between the standard labor supply model our experiment is built around and a standard principal-agent framework used in VHBB. For the purposes of our experiment, we simplify this further by ignoring the complications regarding the constant[5] and explain our payoff function to the subjects as

$$\pi_i(e_i, e_{-i}) = f + b * \min(e_i, e_{-i}) + c * l_i \tag{3.3}$$

With this specification of a payoff function we implement a standard version of a coordination game with costly effort by setting $f = 60$, $b = 20$ and $c = 10$. We allow unbounded output, however for ease of exposition and comparison to VHBB, if the total number of tasks were capped at 7, as in VHBB, the payoffs in the costly treatment would mimic those in Table 3.1.[6] For the costless effort case, we set $c = 0$ and obtain the matrix in Table 3.2.

In the costly effort case, we clearly see the standard coordination game structure in

[5]We do this to keep the overall payment level the same between the costly and costless effort treatments. Using the strict version of the payoff function, the fixed payoff would differ between the two treatments due to the $c * T$ term going to 0. We will instead hold the fixed payment common between both treatments. This does not impact the marginal incentives.

[6]VHBB had integer effort choices ranging from 1 to 7. We also include the option of 0 in our examples to the subjects, so we have presented this version.

		Minimum of the Team's Contribution							
		7	6	5	4	3	2	1	0
Own	7	200	180	160	140	120	100	80	60
Contribution	6		180	160	140	120	100	80	60
	5			160	140	120	100	80	60
	4				140	120	100	80	60
	3					120	100	80	60
	2						100	80	60
	1							80	60
	0								60

Table 3.2: Coordination game without effort cost.

which group members choosing the same number is an equilibrium with the Pareto dominant case at the maximum contribution. Coordination may be difficult of course because contributing above the team minimum is costly. In our real effort version, coordination may be particularly difficult because there is not a common upper bound to coordinate on due to the fact that subjects will be heterogeneous in their ability. In the costless effort case, there is no cost of contributing effort above the team minimum and the Pareto dominant solution is still the maximal contribution to the group account. These structures mimic those used in the stylized experiments of VHBB.

As mentioned, in VHBB total production is essentially capped at 7. In order to better mimic how real effort experiments are typically conducted, we impose no such cap and allow subjects to complete as many instances of the counting task as they like in each two-minute round. They receive base earnings of \$0.60 per round and then receive \$0.20 times the minimum of the number of these tasks completed by any member of their 4 person group. In the costless effort treatment, subjects are only allowed to complete tasks on behalf of the group. Consequently, completing more tasks than the minimum of others in their group imposes no opportunity cost on them. In the costly effort treatment, subjects are able to complete instances of the counting task on behalf of the group yielding earnings of \$0.20 for the minimum number completed by all members of the group or they can choose to complete instances of this same task to generate earnings only for themselves. Each task completed on behalf of themselves earns \$0.10. Consequently, any instances of the task that an individual completes for the team above the minimum of others leads to a payment of \$0.00 and a foregone payment of \$0.10 they could have received had they directed the output to the individual account. Thus an opportunity cost is imposed and effort is costly. On the total productivity range of 0-7, Tables 3.1 and 3.2 accurately reflect the relative incentives. Even though our design allows subjects to complete more than 7 instances of the task (and they do), the same pattern continues to higher levels of productivity. They engage in this game for 10 periods with the same four person group and feedback was given after each round.

Examining Tables 3.1 and 3.2 shows that there are two real differences between the treatments. The first is the intended one of the treatments differing as to whether contributing above the team minimum was costly or not. The second is an unintended consequence of disallowing any earnings from the outside option which leads to coordination earnings for the 0-7 productivity example to be in the range \$0.60 to \$2.00 in the costless effort case but \$1.30 to \$2.00 in the costly effort case. VHBB gets around this problem by imposing a rule that essentially makes effort costly up to the group minimum but not above. This rule would have been difficult to implement in our real effort setting. This difference suggests that the gains from coordination are greater in the costless effort case than the costly effort case. We discuss the potential consequences of this element as we present the results.

Experiments 1 (the real effort coordination game) and 2, to be explained later, were conducted in common sessions where Experiment 2 preceded Experiment 1. Both experiments were comprised of two treatments and we conducted two sessions for each treatment. We altered the mix of treatments that subjects observed to control for order effects. For instance, if the first session had treatment A from Experiment 1 and treatment A from Experiment 2, then a different session which also started with treatment A from Experiment 1 would be followed by treatment B from Experiment 2. All subjects were students at Ohio University and the experiment was programmed using Z-tree software, Fischbacher (2007). Table 5.1 provides information on the average earnings and the number of subjects who participated in each treatment order, given as AA, AB, BA and BB. Sessions lasted approximately an hour. Sample instructions for both experiments are given in the Appendix.

	AA	AB	BA	BB
Average Earnings (USD)	\$21.80	\$32.47	\$28.09	\$32.57
Number of Subjects	12	12	12	12

Table 3.3: Earnings and number of subjects by treatment for Experiments 1 and 2.

3.1. Experiment 1 Results

The core question to be addressed in this first experiment is whether providing an opportunity cost for contributing to a group account in a weak link coordination game has the

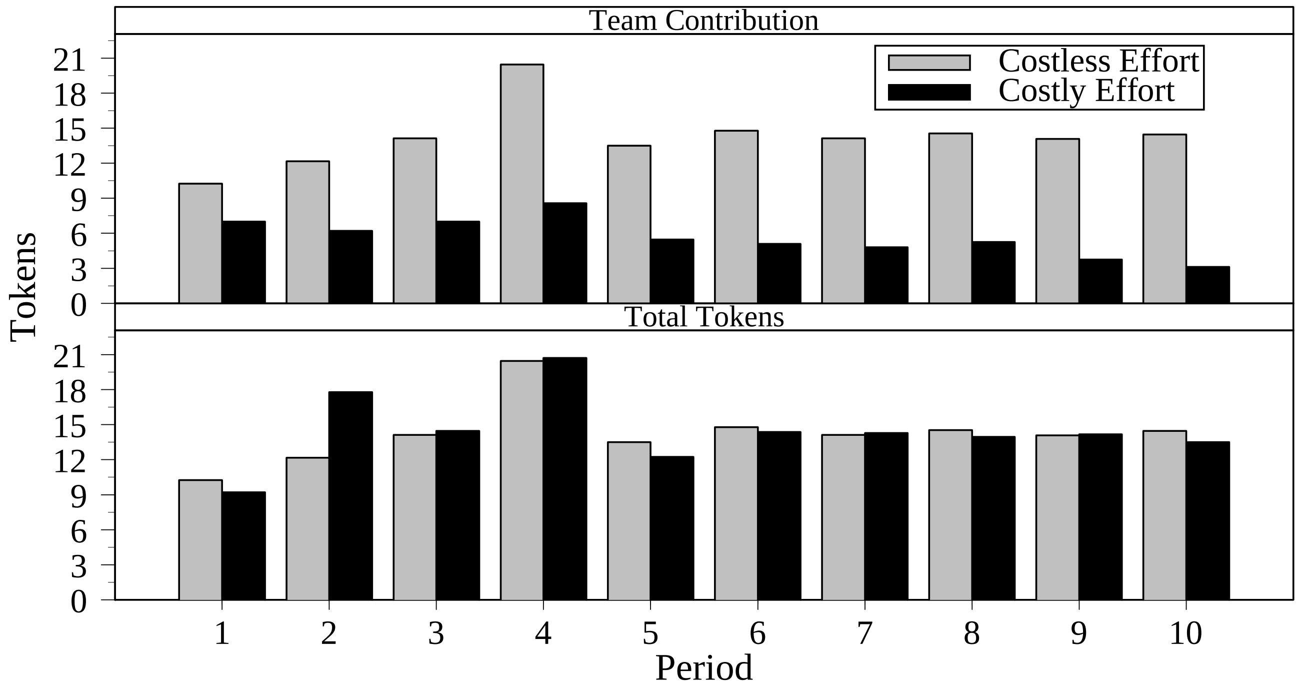

Figure 3.1: Paired bar graphs by period of contributions to the group and total production for both costless and costly effort coordination games.

same impact on coordination in a real effort study as in the stylized effort experiments in VHBB. Figure 3.1 shows the average contribution of tokens to the team account in all 10 periods separated by treatment. It also contains the total production, which for the Costless Effort treatment is exactly the same as the contributions to the team account while for the Costly Effort treatment it includes the contributions to the private account. The basic results are readily apparent. The total team contributions are very high in the Costless Effort treatment, with an average of 14.15, while they are much lower in the Costly Effort Treatment, with an average 5.5. On the other hand, total production appears to be approximately the same between both treatments, with averages of 14.15 and 13.77 respectively. This means that in either treatment, subjects completed essentially the same total instances of the task but in the Costly Effort case, many of these instances are completed for the individual account despite the fact that completing them for the team account would be payoff dominant should all team members choose to do the same.

Table 3.4 provides the statistical analysis to support the visual results. We provide regressions examining how team contributions, the minimum of the team contributions, and total contributions vary by treatment. We also provide specifications examining whether there is a time trend. These regressions are all random effects panel regressions with the standard errors clustered at the team level. In the case of the minimum of team contributions, the observations are at the team level, whereas for team contributions and total production, the observations are at the individual level.

Result 1. *In a real effort design, costless effort yields high coordination while costly effort leads to a breakdown of coordination.*

As predicted, we find that for team contributions and the minimum of team contributions, the treatment effect of Costly Effort is large and highly significant. We also find a time trend result where the individual contributions to the team and therefore the minimum contributions to the team are declining in the Costly Effort treatment, yet they are constant or rising in the Costless Effort treatment. Further we find no difference in the base treatment effect or in the time trend when we look at total production meaning that these effects can be eliminated as potential confounds. Thus, we find that just as with the stylized effort case in VHBB, real effort coordination games demonstrate the same properties that coordination largely fails when effort is costly yet "succeeds" when effort is not costly.

As we noted previously, our treatment also leads to an unintended difference which is that the gains from moving from low to high coordination are greater in the costless effort treatment than the costly effort treatment. One could mount an argument that this provides greater incentives to coordinate on high levels and this alone could explain the higher level of coordination. We will provide an explanation for why this is not a totally compelling claim, but it is also not all that relevant for our core question in this section. Our question is whether or not real effort experiments, by their nature of involving real effort, lead to subjects being able to coordinate as was suggested in Bortolotti, Devetag, and Ortmann (2009). Our results show that this is clearly not the case. Our findings demonstrate that when effort is made costly analogous to the stylized weak link games in VHBB, coordination fails with a real effort design just as it did in the stylized design. Thus one should not conclude that the differences found are due primarily to the differences in effort provision but rather due to differences in the cost function.

When comparing our costless and costly effort treatments, it is possible to argue that subjects were better able to coordinate in the costless effort case due to the higher marginal returns from that coordination. This is not the best interpretation of the results. One way of demonstrating this point is to examine the coefficients on the time variables in Table 3.4, as these variables indicate if information on choices of others are incorporated into

	Team Contributions		Min of Team Contributions		Total Production	
	(1)	(2)	(3)	(4)	(5)	(6)
Costly Effort	-8.621***	-7.250***	-7.783***	-6.333***	-0.379	-0.617
	(0.689)	(0.713)	(1.056)	(1.147)	(0.726)	(0.737)
Last Half		0.1000		0.767**		0.1000
		(0.134)		(0.317)		(0.134)
Last Half * Costly Effort		-2.742***		-2.900***		0.475
		(0.723)		(0.662)		(0.293)
Constant	14.15***	14.10***	10.85***	10.47***	14.15***	14.10***
	(0.454)	(0.472)	(0.753)	(0.870)	(0.454)	(0.472)
Obs (Groups)	480 (12)	480 (12)	120 (12)	120 (12)	480 (12)	480 (12)

Standard errors clustered at team in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 3.4: Random effects panel regressions on team and total production.

subsequent decisions. The results show a very strong negative time trend of individual contributions to the team account in the Costly Effort treatment. What this indicates is that subjects are likely responding negatively to problems coordinating in those sessions by shifting their efforts to the account that generates certain returns. It basically indicates the expected response to mis-coordination which is that people shift away from the group account over time as those contributing above the minimum level of others pull back their effort levels. In the Costless Effort treatment there is no negative time trend on contributions to the team account. The indication is that individuals start off producing as much as they can and do not pull back their contributions even when they realize they have contributed more than others. Why would they pull back? Contributions to the group are costless to them and sitting idly may generate negative utility. This difference in response to mis-coordination is better explained by the differences in the cost of contributing effort above the minimum of others, and these results match with the exact same differences in behavior found in VHBB for their costly and costless effort treatments. Note too that if effort in our Costly Effort treatment is costly due to some notion of "mental costs" or some other notion commonly thought to represent effort costs in real effort designs with no outside option, contributing above the minimum level of others would have likely led to the same decline in contributions over time as in the costly effort treatment. The indication from our results is that whatever those other costs are, they are not significant enough to lead subjects to pull back their contributions when they find that they have contributed above the minimum of others, which is exactly how they respond when doing so is actually costly.

4. Experiment 2 Design: Incorporating Effort Costs Into Real Effort Design

The results of Experiment 1 make it clear that the proper specification of an outside option is important for real effort experiments so the experimenter can draw correct inferences on the cause of observed behavior. That leads to the question of how one could bring effort costs into a standard real effort design. In this section we propose and then test the effectiveness of an effort cost inducing module that could be added to any real effort experiment to allow effort costs to be included and controlled for. The fundamental innovation is the addition of an outside option that implements an effort cost function as specified in the standard model. We will add this outside option to the experiment design of Araujo, Carbone, Conell-Price, Dunietz, Jaroszewicz, Landsman, Lamé, Vesterlund, Wang, and Wilson (2016) and attempt to determine if we can recover the theoretically predicted wage effect once a cost of effort is included in the experiment.

There are several prior papers that add outside options but most of these prior implementations have some drawbacks which keep them from being universally useful. A commonly thought of and occasionally used outside option is to simply allow subjects to browse the internet (e.g., Corgnet, Hernán-González, and Schniter (2015)). While this task has relevance to many external situations, implementation in the laboratory is problematic. This is primarily due to the time structure of most experiments. Most designs have production periods which may be only a few minutes. Subjects switching between internet browsing and the experiment involves relatively substantial switching costs and subjects may not find 30 seconds of internet browsing valuable, especially using an unfamiliar browser lacking their normal bookmarks. This is despite the fact that workers in office jobs may indeed choose to spend hours online rather than engaging in their primary work activity. The other important drawback of methods like this is that the experimenter has neither knowledge of nor control over the value of these activities to their subjects. This means that for some subjects giving up 30 seconds of time browsing the internet is very costly but not others, and this idiosyncratic difference which may have little relevance to any treatment condition could drive treatment differences. In sum, it is unclear that giving up short periods of time browsing the internet is costly and the lack of control over this cost could be a confound of a treatment effect.

There are also prior papers which provide outside options that have fixed or linear utility structures (e.g., Johnson and Salmon (2016)). While these designs can be effective in some cases, those payoff structures are not sufficient to guarantee an interior optimum for the subjects. That is, if the "effort" activity pays back at some piece rate and the leisure option pays back at some other rate, the effort activity could still dominate the alternative through the entire production period. While this demonstrates a revealed preference relationship that working is revealed preferred to this alternative option, it can be difficult to observe treatment effects in some cases due to the boundary solutions.

What is needed is an outside option with a payoff structure that is non-linear, as assumed in the standard theoretical models of effort provision, such that it is possible to expect interior optima for effort provision. The design should also fix the cost of effort to be the same between all subjects. Our design satisfies these needs. In the Araujo et al. experiment, subjects are engaging in a standard slider alignment task based on Gill and Prowse (2011) in which they are paid a piece rate wage for each slider aligned. In our version subjects have 3 minutes to align as many sliders as they wish facing either a \$0.01/slider or \$0.04/slider wage rate. After subjects have completed an initial round of this task with no outside option, we then introduce the option to engage in an alternative to the slider task. The alternative activities we provide are playing Tic-Tac-Toe (TTT) against a mildly challenging computer algorithm and solving word search matrices where the subjects find words embedded into matrices of letters. The nature of these tasks is not important, they merely need to be active and perhaps mildly amusing for the subjects. The key is in how these alternative activities are incentivized.

When an outside option is available, subjects are allowed to spend time aligning sliders or they can switch to a screen with the outside option tasks. Their earnings are based on how many sliders they complete (piece-rate wage) and the total time spent on the outside option screen. To make the experiment easier, they begin the production round facing the slider screen and are allowed to switch to the outside option screen whenever they like but the switch is only allowed once. Thus once they switch away from aligning sliders, they cannot switch back. This is not a necessary element of the design, but chosen to allow for a cleaner design.[7] If a subject chooses to spend the entire 3 minutes on the outside option screen, they earn a fixed amount which we set at \$1.19. This fixed payment specifically does not depend on how many TTT games they win or how many words they find in matrices. Any amount of time subjects choose to spend aligning sliders before switching over to the outside option screen decreases this amount. The total cost of any amount of time spent aligning sliders is $0.006t^{2.3}$ where t represents seconds spent aligning sliders. This generates a convex time cost function, as is assumed in standard models. Note that this structure exactly matches with the specification in Equation 2.4.

[7]For an experiment with longer production periods allowing frequent switching would be reasonable and easy to implement. For our simple experiment here it seemed an unnecessary complication to explain to subjects.

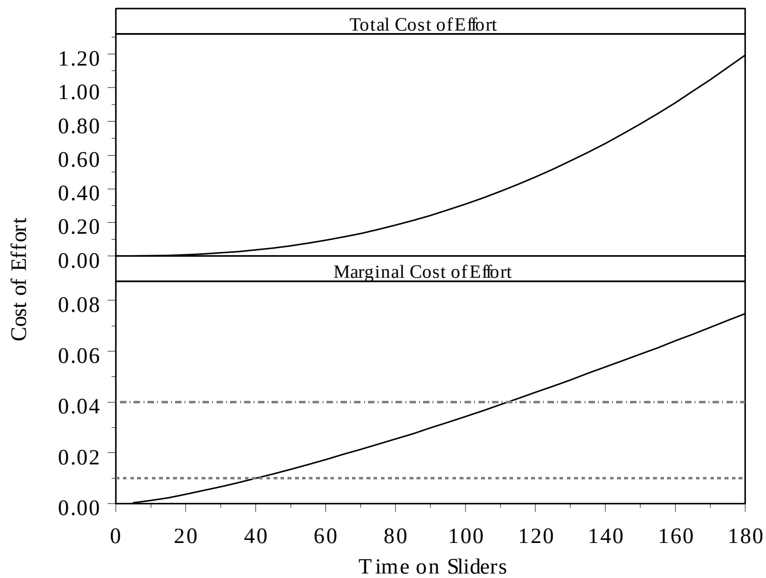

Figure 4.1: Total cost of effort and then Marginal Cost for 5 second intervals.

To make the implications of this cost function easier to understand, we do not present the mathematical version of this function rather we alert them on their screen to how much their earnings will decrease for the next five seconds they spend aligning sliders. We chose that time increment because in the Araujo et al. data, subjects on average aligned 1 slider every 5 seconds. We also have the subjects complete one initial period of slider alignment without the outside option. From the initial period performance, we calculate how many sliders they align on average in 5 seconds and provide this information to them during the production periods with the outside option. This makes it easy for a subject to determine the point at which they believe aligning sliders is no longer worthwhile for them. Figure 4.1 shows the total cost function and the marginal cost of effort for 5 second intervals. This construction allows for a way to predict when a subject will stop working and will choose the outside option if they are sensitive to incentives in a real effort task. At the \$0.01/slider wage rate, an individual who completes on average 1 slider per 5 seconds will find that time spent aligning sliders is more valuable than their outside option for 40 seconds. After that, the marginal cost of foregoing the outside option dominates their earning ability in the slider alignment task. At the \$0.04/slider wage rate, aligning sliders is more lucrative than the alternative up to 110 seconds. If an individual is faster or slower at aligning sliders, then their optimal time to spend aligning sliders will shift accordingly but it should still be case that the switchover point should move up with a wage increase.

This design is a careful implementation of the standard principal agent model in that subjects have an outside option making effort costly and we can control the nature of that cost. For other experiments, one could easily implement different cost functions, heterogeneous cost functions or any other elements thought to be important. Simpler implementations of the time increments would also be easy to achieve if that were a better fit with an experiment. We conduct experiments with this design as a simple test of whether once effort costs are present, does a wage effect emerge with subjects choosing to align more sliders as the wage rises. If this holds, then it serves as a simple demonstration of how an outside option of this nature can be added to any experiment to restore theoretical credibility to real effort designs.

Experimental procedures and other relevant information was already described above. Sample instructions are given in the Appendix.

4.1. Experiment 2 Results

The issue we wish to examine with this experiment is simply whether the introduction of an outside option allows us to observe a wage effect and whether this effect is predictable using a standard model. Figure 4.2 provides a visual representation of the data. It shows the number of sliders aligned in periods 2, 3, and 4 of the experiment which are the periods where the outside option screen is available (recall period 1 does not include an outside option and is used as a baseline measure of performance/ability in the task). The immediate observation from this figure is that subjects aligned more sliders in the High Piece Rate treatment than in the Low Piece Rate treatment. We also provide lines indicating the number of sliders that on average would be optimal for subjects to complete based on their speed of slider alignment in period 1. Not surprisingly, the prediction is higher in the High Piece rate treatment as the benefit for aligning sliders is higher. We note that our subjects exceed this prediction in all rounds. This is in part due to a learning effect as subjects were able to complete sliders more quickly in periods 2-4 than in period 1. Given that, they should be predicted to do more than indicated by those lines.[8]

Table 4.1 provides the statistical analysis of the slider data. We provide the analysis using three different regression approaches and two specifications for each. Columns 1 and 2 present OLS random effects regressions with the standard errors clustered at the subject level. Columns 3 and 4 provide random effect Tobit regressions to correct for the

[8]We could have instead demonstrated the wage impact by showing that subjects spend more time aligning sliders with a higher wage as they spend 86 seconds on average aligning sliders in the High PR treatment but only 50 seconds in the Low PR treatment. As the two measures are almost perfectly collinear, we chose to go with sliders completed as the more typical metric.

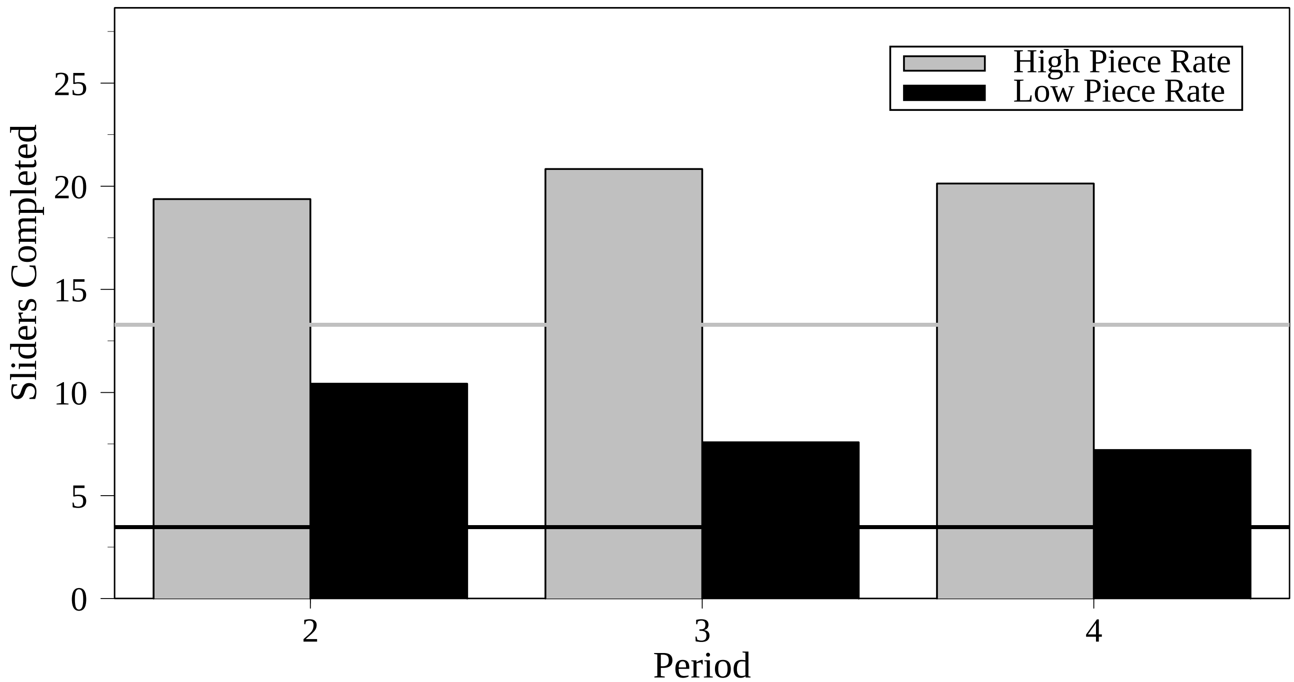

Figure 4.2: Number of sliders aligned by treatment in the main production periods. Light gray line is the predicted number of sliders based on round 1 speed for the Low Piece Rate treatment. Black line is prediction for the High Piece Rate treatment

fact that a number of observations are piled up at 0, especially in the Low PR treatment. Columns 5 and 6 provide standard Tobit regressions with the standard errors clustered at the individual level. For each approach we provide one specification to examine the base treatment effect and then a second to examine the data for any time trends as well as whether the time trend depends on the treatment. In all cases we include the individuals' period 1 productivity as a measure of their ability.

Result 2. *Higher piece rate wages increase effort in the presence of an outside option.*

The results are consistent across all specifications. We find that the treatment variable for the High PR treatment is large and highly significant indicating that a wage effect is clearly present in the data. We also find there to be no time trend for either treatment. Period 1 productivity is highly correlated with later period performance, as would be expected, as this is a measure of their speed of slider alignment and the more they complete in period 1, the more they should be able to profitably complete in the periods with the

	RE with CSE		Panel Tobit		Tobit with CSE	
	(1)	(2)	(3)	(4)	(5)	(6)
High PR	9.167***	6.417*	14.14***	10.71**	13.59***	10.20**
	(2.883)	(3.848)	(3.782)	(4.458)	(3.741)	(4.536)
Period 3		-7.125		-9.228		-9.996
		(7.670)		(6.886)		(10.81)
Period 4		-7.167		-9.537		-8.490
		(6.155)		(6.918)		(7.344)
Period 3 X High PR		4.292		5.343		5.727
		(4.009)		(4.180)		(5.527)
Period 4 X High PR		3.958		5.004		4.458
		(3.312)		(4.197)		(3.882)
Period 1 Total	0.622***	0.622***	0.733***	0.733***	0.716***	0.715***
	(0.179)	(0.181)	(0.210)	(0.211)	(0.197)	(0.197)
Constant	-14.43***	-9.668	-27.36***	-21.13***	-25.82***	-19.66**
	(5.154)	(7.479)	(7.219)	(8.169)	(7.447)	(9.567)
Obs (Subjects)	144 (48)	144 (48)	144 (48)	144 (48)	144 (48)	144 (48)

Cols (1),(2),(5) and (6) have standard errors clustered at the individual level. *** p<0.01, ** p<0.05, * p<0.1

Table 4.1: Random effects panel regressions of sliders completed.

outside option.[9] These results allow us to refine the conclusions drawn in Araujo et al. regarding the reliability of the slider task. In that paper, the authors suggest that the lack of a wage effect in their experiments is because the slider task may not be an effective or useful task to use in laboratory experiments. Our analysis shows that there is nothing wrong with the slider task itself, but there is a limitation in how it is typically implemented. It is the lack of a viable outside option that creates the lack of a wage effect and adding in an outside option using a simple method such as we have done here makes the slider task capable of serving its purpose. The same will be true of many other tasks used in real effort experiments as the general issues we examine here should not be expected to relate only to the slider task.

[9]We note that one reason we need to include the period 1 sliders is that there was substantial heterogeneity in period 1 productivity. In the Low PR treatment, subjects completed on average 22 sliders while they completed on average 26 in the High PR case. This difference is not significant at the 10% level for either a standard t-test or a Wilcoxon rank-sum test. Not only is this difference not as large as the difference once the outside option is allowed, but with this initial (insignificant) difference controlled for, it ensures that the treatment effect we find is due to the treatment and not any heterogeneity in speed across treatments.

5. Experiment 3 Design: Is Real Effort Just Different?

The first two experiments demonstrate the importance of effort costs in real effort experiments and provide ways to induce them. In this third experiment, our focus shifts to testing a claim often made in the literature which is that real effort is somehow just different than stylized effort and it should be expected to yield different results. As discussed before, there have been papers demonstrating a divergence in results between stylized and real effort experiments but largely these cases have examined two environments which are not directly comparable due to differing effort costs between effort modalities. Our goal in this section is to test the two approaches to modeling effort in a setting in which we control for as many differences as possible between real and stylized designs to determine if other elements can still drive a difference. We do this in a setting in which there is still an *a priori* reason why one might expect the two forms of effort to lead to different results, which we hope gives real effort a clear chance to generate different behavior than stylized.

We have chosen to implement this test in the context of a standard public goods experiment matching the basic design stemming from Isaac, Walker, and Thomas (1984) and Isaac and Walker (1988). One benefit of using this design is that, as explained in Ledyard (1995) and Chaudhuri (2011), the base results have been replicated many times allowing us ample comparisons with prior work. We examine real and stylized versions of this design and all treatments possess the same incentive structure. Participants are randomly assigned to a group of four and remain matched for the duration of the experiment. In each period, individual subjects accumulate tokens over time either by their own effort or by a random arrival process and can choose to direct the tokens towards a group account or an individual account. For each token invested in the individual account, the participant earns \$0.20. For each token invested in the group account, the group earns \$0.40 with group earnings divided equally among 4 group members. This leads to a marginal per capita return (MPCR) of 0.5 and sets up the standard incentive conflict known as a social dilemma. At the end of each period, participants are provided with feedback that includes a reminder of their contributions, the total number of tokens contributed to the group by all members of the group, and a summary of their earnings for the period. This same process is repeated for ten periods.[10]

An important aspect of our design is the need to keep the maximum possible effort

[10] There are other papers that have real effort in VCM's, for example Van Dijk, Sonnemans, and Van Winden (2001) had subjects solve two-variable optimization problems while Cooper and Saral (2013) use GMAT questions.

common across all subjects and common across effort modes. In many real effort designs, maximum effort or productivity is unbounded and therefore subject specific. If we allowed this in our design, it would induce heterogeneity in endowments in the real effort treatments while the stylized effort treatment would exclude this feature. The literature utilizing stylized designs already show that heterogeneity will affect behavior (e.g. Cherry, Krol, and Shogren (2005); Buckley and Croson (2006); Reuben and Riedl (2013)) and so this would be a substantial confound in comparing behavior between treatments. We fix the token budget at 10 regardless of the treatment.[11] A fixed budget like this is the same as limiting subjects to the same number of hours in a day, which is perfectly reasonable, though we are also then insisting that the production functions of all subjects are the same. While the latter part of that may not be generally true, implementing the experiment this way allows for a much cleaner comparison.

Another potential source of behavioral differences between real and stylized effort designs is the amount of time involved in decision making. In the standard VCM, a subject must make a single decision about token allocation and periods can go very quickly. In the real effort version, subjects spend time on the real effort task producing their tokens and their contributions to the accounts. The timing difference could, for instance, lead to a person becoming either more or less thoughtful over their contribution choices, which could lead them to be either more or less cooperative. This element must also be eliminated as a difference between treatments.

Real effort tasks may also differ from stylized effort due to potential differences in cognitive load. Engaging in the real effort task could distract participants from the VCM task triggering different behavior. While the directional impact of such a distraction is unclear, it seems quite clear that contribution decisions could be impacted. Our design must therefore equalize cognitive load between treatments to the extent possible.

The treatments were designed to address each of these concerns. One further element we examined is that many prior papers on real effort make a distinction between what we will term Useful Effort and Trivial Effort. In the former, the effort is on a task that seems useful and could benefit someone, e.g. stuffing envelopes with mailers for an academic department or shelling walnuts for a bakery. In the case of Trivial Effort, the effort is clearly just an artifact of the experiment and being conducted for no other purpose, e.g.

[11] We did not impose this restriction in the coordination game experiments (Experiment 1) presented earlier. The reason for this is that in the first experiment our goal was not to make a tight test comparing the specific choices made in real versus stylized effort, rather we wanted to examine whether we would observe the same comparative static response. In this experiment we will be comparing effort choices directly to see whether one mode leads to greater effort than the other. Thus the need to keep maximum productivity the same between the two effort styles in this design.

Fund Name	Ticker	NAV as of 10/5/2015	Total Net Assets	Load Adjusted Returns			
				1 Yr Return	5 Yr Return	10 Yr Return	Since Inception
Advisory Rsrch Glbl Val	ADVWX	11.47	$14,700,000	-7.28%	9.11%	N/A	9.38%
AllianBer Gl Value;A	ABAGX	N/A	N/A	0.90%	6.26%	2.35%	2.59%
AllianBer Gl Value;Adv	ABGYX	N/A	N/A	5.61%	7.49%	3.09%	3.33%
AllianBer Gl Value;B	ABBGX	N/A	N/A	0.70%	6.36%	2.02%	2.29%
AllianBer Gl Value;C	ABCGX	N/A	N/A	3.67%	6.42%	2.06%	2.32%

Figure 5.1: Sample of data subjects would enter in the Useful Effort treatment.

aligning sliders. We conduct our real effort design under both structures to test if we can detect a difference in these two environments. In the Useful Effort (UE) treatment, subjects engaged in a data entry task in which they enter actual financial data from Reuters into a database. This data is an important component of a research project of another faculty member at the university where the experiments were conducted (not a coauthor on this project). In the instructions we explain to the subjects very clearly that the data entry task is vital to the research of this faculty member and exhort them to be careful in their work. This was an attempt to have subjects truly see this as useful effort and not an abstract task necessary only for the experiment.

Subjects earn a token by entering a line of data from a sheet provided to them. An example of this data is shown in Figure 5.1. Each line of data earns a single token and all subjects are required to enter exactly ten lines per period. While earning the tokens, subjects choose how to allocate their productive time between two utility earning activities, exactly as in the labor supply model outlined above. They do this by using a toggle button on their screen. At the beginning of a period they must switch the toggle to either the group or private account. After they make an initial choice they begin producing tokens and any tokens they produce go to the selected account. They can switch the toggle as many times as they like and at any time, with tokens accruing to whichever account is active when they submit a line of data. Subjects can therefore choose how much of their time/effort to devote towards working for the public account or working for themselves.

The Trivial Effort (TE) treatment is conducted identically in all aspects to the UE treatment except that the data subjects enter is presented to them with no context. Subjects in the TE treatment are handed identical data sheets to those in the UE treatment but there was no mention that the data would be used for any external purpose. They are only told that the reason to enter the data is to earn the tokens. To accomplish this, two copies of each data sheet were printed where one went to someone in the TE treatment and one went to someone in the UE treatment. The data generated by the TE treatment was discarded while the data generated by the UE treatment was passed along to the researcher who needed it.

Designing a treatment to represent Stylized Effort (SE), which has the properties of a standard VCM but that differs from the previous two treatments only by how the tokens are earned, requires the design to be somewhat different from a standard VCM. In this treatment, subjects receive tokens but do not engage in the data entry task. In order to ensure that the timing issues are the same between this treatment and the others, subjects make their investment choices using the same toggle switch as in the other treatments but instead of earning the tokens through data entry, the tokens arrive at random intervals. This means that they are making a stylized effort choice as they are choosing how to allocate effort/tokens towards each account but no actual effort is required from them. The token arrival times are drawn at random from the actual distribution of subject times to complete a data entry line from the UE and TE sessions. The average length of time between tokens is 22 seconds, with a maximum of 73 seconds and a minimum of 8. Before the token appears, subjects receive a warning that a token is going to be deposited into the selected account in three seconds which allows them time to change the current account to which tokens are accruing. In order to provide a similar cognitive load to the data entry task, subjects are able to play Tic-Tac-Toe against a computer opponent for no earnings while the tokens are arriving. It is made clear that playing this game is not connected to earning tokens and that there are no earnings related to playing.[12]

The underlying incentives for all treatments are a precise analog of the labor supply model shown in Equation 2.1. Each individual has a fixed budget of tokens (time) to split between two utility providing alternatives. In our context, an individual chooses how much effort to supply to the group account (labor) with the rest of their effort allocated to the private account (leisure). The value of this structure is that the cost of each token/unit of effort contributed to the group account is the private account earnings foregone. This is true for all treatments. Of course one might still be concerned that we aren't capturing all of the mental or additional effort costs we described in the theory section. While we remain skeptical that those costs are substantial, and our first two experiments support our skepticism, subjects in this phase all have to enter 10 lines of data regardless of how they divide this up. Hence, whatever mental costs are for entering the data, they are held constant regardless of the division of tokens and should not enter to the decision regarding token division.

This third experiment is intentionally designed to provide a channel through which

[12] 66% of subjects played at least one game of tic-tac-toe. The average number of games played in a period was 11.

the mode of effort could affect decisions. The base incentives make it a dominant strategy for all effort to be allocated to private production, but we know from many prior public goods studies that this does not usually occur. People generally engage in conditional cooperation in which they contribute to the public good so long as others do, but we also generally observe that people's willingness to contribute declines over time. There is also literature demonstrating that subjects often feel a greater sense of entitlement to earned money than to money received at random, see for example Hoffman, McCabe, Shachat, and Smith (1994). A plausible hypothesis is that compared to someone who has been gifted with tokens, a person who must work for their tokens might experience greater disutility from finding that they have contributed more to the group account than other group members. If valid, then the real effort treatments would lead to different behavior than the stylized effort treatment as subjects may be less willing to risk cooperation with others or their contributions might decline faster if they are disappointed with the contributions of their group members. The need to allow for such a channel is why the public goods framework was chosen.[13] There are certainly other such channels through which real effort could differ from stylized effort, and we make no claims here of comprehensive testing.

We conducted 2 sessions of all treatments. All subjects were students at Ohio University and the experiment was programmed using Z-tree software, Fischbacher (2007). Table 5.1 provides information on the average earnings and the number of subjects who participated in each treatment. Sample instructions are given in the Appendix.

	Useful Effort	Trivial Effort	Stylized Effort
Average Earnings (USD)	\$30.37	\$31.84	\$30.91
Number of Subjects	28	32	28

Table 5.1: Earnings and number of subjects by treatment.

5.1. Experiment 3 Results

A first look at the data is provided in Figure 5.2 which shows average contribution levels into the group account by period for all three treatments. Since the SE treatment is different from previous stylized effort VCM designs we also include data from two prior

[13] The design seems similar to one that would test whether earned endowments in public good games lead to different behavior than unearned endowments. This has been previously investigated in the public goods framework (e.g. Cherry, Krol, and Shogren (2005), Clark (2002)) with the general finding that earned endowments do not change contributions to the public good. This was not an intentional part of the design but we note the similarity.

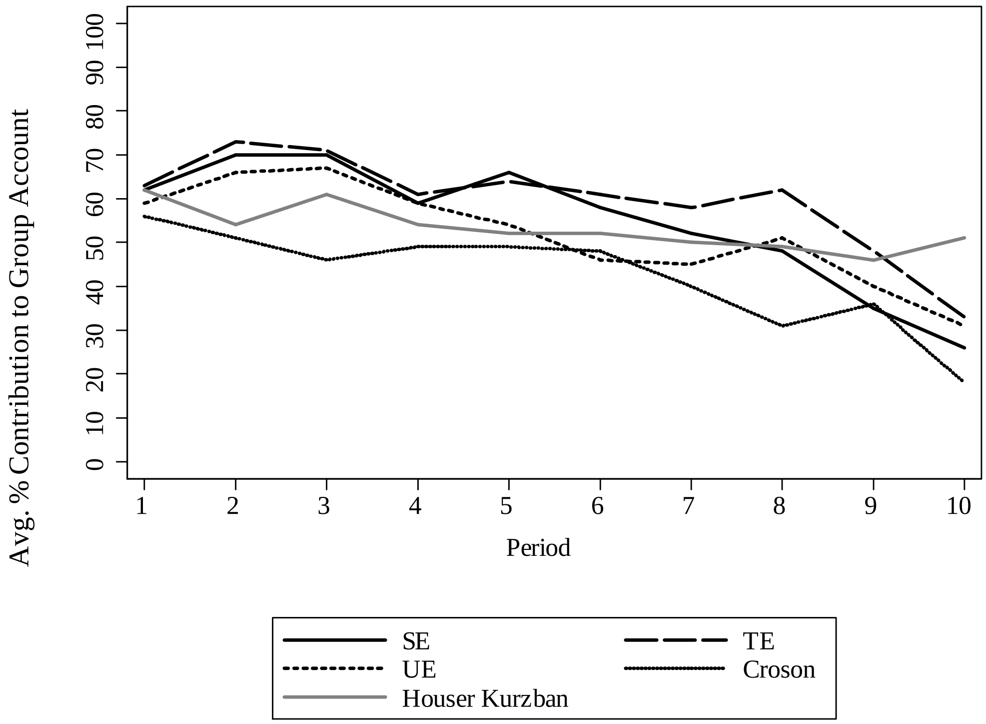

Figure 5.2: Average contribution to the group account by period over all 10 rounds.

studies, Croson (2001) and Houser and Kurzban (2002), which use the more traditional design and also have the same parameterization as our design in regards to group size and MPCR, though they obviously use a different subject pool. The figure shows that the results from all three of our treatments are very similar to each other and the behavior matches what was observed in the prior studies using the traditional stylized design. Specifically, contributions to the public good start at around 60% and steadily decline over time. This similarity suggests that some of the confounds we worried about in the design of our SE treatment may not have been empirically substantial but that is of course only knowable once the data have been gathered.

Table 5.2 summarizes the average contributions to the group account and their standard deviation by period for our three treatments. The table provides additional evidence of the similarity between treatments; the average levels of contributions are similar between treatments as are the standard errors.

Initial tests of these aggregate statistics find no significant differences between any treatments ($p > 0.42$ for all pairwise comparisons using Wilcoxon tests on the average contribution to the group account for each group over all 10 periods); simple $t-$tests on these same group averages also support this result ($p > 0.49$ for all pairwise compar-

Trt	Pd 1	Pd 2	Pd 3	Pd 4	Pd 5	Pd 6	Pd 7	Pd 8	Pd 9	Pd 10	Overall
SE	6.21	6.96	6.96	5.86	6.61	5.75	5.21	4.79	3.54	2.64	5.45
	(2.83)	(2.87)	(3.20)	(3.95)	(3.50)	(3.84)	(3.90)	(3.97)	(3.45)	(3.47)	(3.73)
TE	6.25	7.31	7.06	6.13	6.38	6.09	5.78	6.19	4.75	3.25	5.92
	(2.71)	(2.79)	(3.23)	(3.47)	(2.93)	(3.41)	(3.63)	(3.19)	(3.79)	(3.37)	(3.41)
UE	5.92	6.64	6.68	5.86	5.39	4.64	4.50	5.07	4.04	3.11	5.19
	(3.52)	(3.42)	(3.40)	(3.58)	(3.55)	(3.61)	(3.38)	(3.44)	(3.85)	(3.34)	(3.62)
All	6.14	6.99	6.91	5.95	6.14	5.52	5.19	5.39	4.14	3.01	5.54
	(3.00)	(3.01)	(3.24)	(3.62)	(3.32)	(3.63)	(3.64)	(3.54)	(3.70)	(3.36)	(3.59)

Table 5.2: Mean and standard deviations of contributions to the group account by treatment (Trt) and period (Pd).

isons). If we conduct these tests on group averages by period, again all differences remain insignificant ($p > 0.30$ for all pairwise comparisons).[14]

The simple distribution tests do not correct for the panel structure of the data nor do they allow us to investigate the conditional nature of the decisions. To account for this, Table 5.3 reports results from random effects regressions where the dependent variable is the amount an individual contributed to the group account, with errors clustered at the subject level.[15]

Result 3. *There are no statistically significant differences in contributions between stylized effort, trivial real effort, and useful real effort treatments.*

Each regression specification includes dummy variables for the TE and UE treatments. The first specification includes only these dummy variables and a constant, which provides a clean test for differences between the overall contribution levels. Neither coefficient is significant which indicates that the contributions to the group account in TE and UE are not significantly different from SE. Since the two coefficients are opposite signs, it could be the case that the average level of contributions to the group account in TE and

[14] We can also do standard power analysis to determine what sample would be required in order to find a significant effect given our sample averages and standard deviations. For the $t-$tests on overall contributions by groups we find that assuming a power level of 80%, for the observed effect size to be found significant at the 5% level between the SE and TE treatments, 474 groups would be needed; between the TE and UE treatments, 236 groups; and between the SE and UE treatments, 1716 groups. Recall that the claims in the literature motivating this experiment are that experiments using real effort are more externally valid than stylized effort experiments. Were that true, we should have seen substantial differences in behavior that would have been detectable even at our sample size. The fact that the small differences we observe would require such large samples to discriminate between treatments suggests that any behavioral differences are at best minor.

[15] In an appendix we provide Tobit regressions to correct for censoring and specifications using the Bias-Reducing Linearization (BRL) procedure (McCaffrey and Bell (2002)) to allow us to also cluster for groups despite our small sample. Qualitative results are unchanged. Conclusions are the same from either method.

	(1)	(2)	(3)
TE	0.465	-0.175	-0.321
	(0.602)	(0.825)	(0.865)
UE	-0.268	-0.712	-0.497
	(0.692)	(0.973)	(1.018)
Period		-0.423***	
		(0.089)	
Period∗TE		0.116	
		(0.112)	
Period∗UE		0.081	
		(0.117)	
Group_{t-1}			0.522***
			(0.121)
Group_{t-1}∗TE			0.096
			(0.149)
Group_{t-1}∗UE			0.076
			(0.198)
Constant	5.454***	7.781***	2.355***
	(0.474)	(0.692)	(0.638)
Obs (Clusters)	880 (88)	880 (88)	792 (88)

Clustered robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 5.3: Random effects panel regressions on contributions to the group account.

UE are different. A post-estimation Wald test yields $p = 0.24$ indicating that those two coefficients are also not significantly different from each other.

We also test whether there are differences over time. Figure 5.2 indicated that the time paths look similar, but to test this observation formally the second specification in Table 5.3 includes a time trend and interactions with the treatments. This yields the next result.

Result 4. *There are no statistically significant differences between our three treatments in regards to how contributions adjust across time.*

In this specification, the coefficient on the Period variable is negative and significant indicating that there is decay in contributions over time in the SE treatment. The interactions between Period and the two other treatments are insignificant and of the same sign indicating that the rate of decay is not different for any of our treatments. Post-estimation Wald tests confirm no difference in the decay (the interaction terms, $p = 0.72$) or the treatments after accounting for the decay (the binary variables for the UE and TE treatments, $p = 0.51$) for the two real-effort treatments.

The last regression specification examines how individuals respond to the average contribution level of their group members from the previous period. Many prior studies conclude that subjects often engage in conditional cooperation in VCMs; they are willing to cooperate if others also cooperate but contributions will decline if they do not see others contributing at their expected level. Arifovic and Ledyard (2012) demonstrate that this behavior can yield the standard decay pattern we observe over time and so this last specification includes a lag of the average contribution by an individual's group members (Group_{t-1}) to determine if we can see any differences between treatments regarding how individuals react to the contribution levels of others.[16] This leads to our last result.

Result 5. *There are no statistically significant differences between our three treatments in how contributions adjust to contributions by group members.*

Consistent with conditional cooperation behavior, the significance of the lagged variable indicates that subjects do adjust their contributions based on the contributions of others. The interaction terms, however, are all insignificant which confirms that there are no differences between treatments in regard to how individuals react to the contributions of others. Post-estimation Wald tests support the lack of a difference between the TE and UE coefficients ($p = 0.86$) or the interaction effect between the two real-effort treatments ($p = 0.91$).

6. Conclusion

It is commonly argued that an experiment design based on a stylized effort task does not generalize to field settings as readily as a real effort experiment because the effort cost in a real effort experiment better represents effort costs outside the laboratory. We demonstrate that this claim is simply unfounded. Theoretically, effort costs are properly understood as opportunity costs and yet in most real effort experiments there are no significant opportunity costs. Further, the nature of effort costs depend crucially on the nature of the alternative activity available to an individual, not the work or effort task, making it clear that most real effort designs do not include effort costs that are at all reflective of field relevant effort costs. Consequently, many of these real effort designs have instead been constructed in such a way as to minimize external validity, according to the logic that some have used to argue in their favor. We present three experiments

[16] Because of the near perfect (negative) correlation between the lagged term and Period ($p < 0.0001$), the regression from column two cannot include both Period and Group_{t-1}.

to examine this issue from several different perspectives. We first examine some prior claims suggesting that real and stylized effort yield different behavior in the context of coordination games as claimed in Van Huyck, Battalio, and Beil (1990). Our results suggest that a better interpretation of those results is that behavior differed between real and stylized treatments due to the fact that effort costs differed markedly between the two treatments and not because of the effort modalities themselves. Our second experiment provides a demonstration of how to include effort costs in a standard real effort experiment. We find that once effort costs are implemented in a real effort experiment, one finds the expected standard wage effects that were lacking from prior studies which did not include relevant outside options. Our last experiment is designed to determine if there is some other intrinsic reason that suggests we should expect behavior in a real effort design to be different from a stylized effort. We construct tightly connected real and stylized versions of an experiment and find that subjects make the exact same decisions regardless of whether their choices are based on real or stylized effort designs.

There are a number of important insights contained in these results, though there are of course important limitations as well. We will discuss the limitations first. While our last experiment showed no differences between real and stylized effort, these results are not sufficient to demonstrate that there cannot be other environments in which differences might exist. Further, we certainly have no intention of advancing such a claim. Also, our results do not indicate or suggest that either real or stylized effort should be always preferred in any experiment. Our view is that there are important differences between real and stylized effort designs which may be important to many research questions. For example, if demographic-based diversity in ability is important to a research question, a real effort design would certainly be indicated. While diversity in ability could be modeled in a stylized design, a real effort design can be used to estimate the degree of naturally occurring diversity and exploit it in a way that a stylized design cannot. It is also possible that there may be some other elements of an individual's utility function regarding their work that a real effort design could better trigger that could be important to a specific research question. On the other hand, our results suggest that for a broad range of questions, there may well be no gain from implementing a real effort design and there are certainly substantial costs from doing so. In these cases, one should not discount the useful insights obtainable from a stylized design.

Our second experiment gives researchers a useful tool to implement an actual effort cost in a real effort design to ensure that the choice environment matches the theoretically relevant one. If one observes individuals choosing to engage in effort in a similar design with actual effort costs, then it establishes a revealed preference relationship indicating that the subjects prefer working over an alternative utility-enhancing option. Many prior real effort experiments cannot claim this finding and it calls into question some of the results found using those designs. In the future, researchers can implement cost functions using this or similar means to avoid this problem. The one we actually implemented is perhaps overly detailed and precise in its construction. Other implementations need not be as precisely done. The key element of the methodology involves paying subjects based on how much time they spend engaging in an outside option which is the same as making subjects pay a cost based on the amount of time they spend on the productive activity. This cost should be deducted from a fixed amount the subject would receive for only engaging in the outside option and adding some convexity to the cost is a good idea. So long as that element is included, any observation of time spent engaging in the effort task yields the revealed preference relationship one needs to establish.

Another implication of our general results is that if a researcher wants to implement effort costs in a real effort design that are truly field relevant, they must allow field relevant outside options and in a manner that people might pursue them in the field. This is something that can be quite difficult to do in a laboratory setting with short production periods. Web based experiments that people complete from remote venues where subjects have access to their "normal" outside options, such as Dutcher (2012), may be able to overcome this limitation. Of course experiments of this nature lead to other limitations and a loss of control that might make them not amenable to all research questions.

The overall point from this analysis is that the choice of real versus stylized effort in an experiment design is both simpler and more complicated than is often considered in prior work. It is simpler in the sense that if implemented properly, for most research questions there should be little expectation of a difference in the results between real and stylized effort based experiments. Thus this choice should not impact the ultimate research findings. The sense in which our work shows that modeling effort is more complex is that we demonstrate that more care needs to be taken in the design of real effort experiments to ensure that there is an actual cost for supplying effort. Without such costs, inference regarding the results may be incorrect. Whichever model a researcher chooses should be tied directly to the main research question in order to best improve the inference obtainable through the experiment. Though there are many claims in the literature that real effort experiments are somehow more insightful or more field relevant, such claims seem unfounded and should largely be ignored when designing or evaluating an experiment.

References

Araujo, F. A., E. Carbone, L. Conell-Price, M. W. Dunietz, A. Jaroszewicz, R. Landsman, D. Lamé, L. Vesterlund, S. W. Wang, and A. J. Wilson (2016). The slider task: an example of restricted inference on incentive effects. *Journal of the Economic Science Association 2*(1), 1–12.

Arifovic, J. and J. Ledyard (2012). Individual evolutionary learning, other-regarding preferences and the coluntary contributions mechanism. *Journal of Public Economics 96*, 808–823.

Bortolotti, S., G. Devetag, and A. Ortmann (2009). Exploring the effects of real effort in a weak-link experiment. Working Paper.

Buckley, E. and R. Croson (2006). Income and wealth heterogeneity in the voluntary provision of linear public goods. *Journal of Public Economics 90*(4), 935–955.

Charness, G., U. Gneezy, and A. Henderson (2018). Experimental methods: Measuring effort in economics experiments. *Journal of Economic Behavior and Organization 149*, 74–87.

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature. *Experimental Economics 14*(1), 47–83.

Cherry, T. L., S. Krol, and J. F. Shogren (2005). The impact of endowment heterogeneity and origin on public good vontributions: Evidence from the lab. *Journal of Economic Behavior & Organization 57*(3), 357–365.

Clark, J. (2002). House money effects in public good experiments. *Experimental Economics 5*(3), 223–231.

Cooper, D. J. and K. J. Saral (2013). Entrepreneurship and team participation: An experimental study. *European Economic Review 59*, 126–140.

Corgnet, B., R. Hernán-González, and E. Schniter (2015). Why real leisure really matters: Incentive effects on real effort in the laboratory. *Experimental Economics 18*(2), 284–301.

Croson, R. T. (2001). Feedback in voluntary contribution mechanisms: An experiment in team production. In M. Isaac (Ed.), *Research in Experimental Economics*, Volume 8, pp. 85 – 97. Emerald Group Publishing Limited.

Dutcher, E. G. (2012). The effects of telecommuting on productivity: An experimental examination. the role of dull and creative tasks. *Journal of Economic Behavior & Organization 84*(1), 355–363.

Engel, R. (2010). Why Work When You Can Shirk? Worker Productivity in an Experimental Setting. *Journal of Applied Business and Economics 11*(2), 104–119. Working Paper Florida State University.

Erkal, N., L. Gangadharan, and B. H. Koh (2018). Monetary and non-monetary incentives in real-effort tournaments. *European Economic Review 101*, 528–545.

Esarey, J., T. C. Salmon, and C. Barrilleaux (2012). What Motivates Political Preferences? Self-Interest, Ideology, and Fairness in a Laboratory Democracy. *Economic Inquiry 50*(3), 604–624.

Fahr, R. and B. Irlenbusch (2000). Fairness as a constraint on trust in reciprocity: Earned property rights in a reciprocal exchange experiment. *Economics Letters 66*(3), 275–282.

Fischbacher, U. (2007). z-Tree: Zurich Toolbox For Readymade Economic Experiments. *Experimental Economics 10*(2), 171–178.

Gill, D. and V. Prowse (2011). A novel computerized real effort task based on sliders. Working Paper.

Hagger, M. S., N. L. D. Chatzisarantis, H. Alberts, C. O. Anggono, C. Batailler, A. R. Birt, R. Brand, M. J. Brandt, G. Brewer, S. Bruyneel, D. P. Calvillo, W. K. Campbell, P. R. Cannon, M. Carlucci, N. P. Carruth, T. Cheung, A. Crowell, D. T. D. D. Ridder, S. Dewitte, M. Elson, J. R. Evans, B. A. Fay, B. M. Fennis, A. Finley, Z. Francis, E. Heise, H. Hoemann, M. Inzlicht, S. L. Koole, L. Koppel, F. Kroese, F. Lange, K. Lau, B. P. Lynch, C. Martijn, H. Merckelbach, N. V. Mills, A. Michirev, A. Miyake, A. E. Mosser, M. Muise, D. Muller, M. Muzi, D. Nalis, R. Nurwanti, H. Otgaar, M. C. Philipp, P. Primoceri, K. Rentzsch, L. Ringos, C. Schlinkert, B. J. Schmeichel, S. F. Schoch, M. Schrama, A. Schütz, A. Stamos, G. Tinghög, J. Ullrich, M. vanDellen, S. Wimbarti, W. Wolff, C. Yusainy, O. Zerhouni, and M. Zwienenberg (2016). A multilab preregistered replication of the ego-depletion effect. *Perspectives on Psychological Science 11*(4), 546–573.

Hoffman, E., K. McCabe, K. Shachat, and V. Smith (1994). Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior 7*(3), 346–380.

Houser, D. and R. Kurzban (2002). Revisiting kindness and confusion in public goods experiments. *American Economics Review 92*(4), 1062–1069.

Isaac, R. M. and J. M. Walker (1988). Group Size Hypotheses of Public Goods Provision: An Experimental Examination. *Quarterly Journal of Economics 103*, 179–199.

Isaac, R. M., J. M. Walker, and S. H. Thomas (1984). Divergent evidence on free riding: An experimental examination of possible explanations. *Public Choice 43*, 113–149.

Johnson, D. and T. C. Salmon (2016, January). Sabotage vs discouragement: Which dominates post promotion tournament behavior? *Southern Economic Journal 82*(3), 673–696.

Ku, H. and T. C. Salmon (2012). The incentive effects of inequality: An experimental investigation. *Southern Economic Journal 79*(1), 46–70.

Ledyard, J. O. (1995). Public Goods: A Survey of Experimental Research. In J. H. Kagel and A. E. Roth (Eds.), *The Handbook of Experimental Economics*, pp. 111–194. Princeton, New Jersey: Princeton University Press.

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995). *Microeconomic Theory*. New York, New York: Oxford University Press.

McCaffrey, D. F. and R. M. Bell (2002). Bias reduction in standard errors for linear regression with multi-stage samples. *Survey Methodology 28*(2), 169–182.

Reuben, E. and A. Riedl (2013). Enforcement of contribution norms in public good games with heterogeneous populations. *Games and Economic Behavior 77*(1), 122–137.

Van Dijk, F., J. Sonnemans, and F. Van Winden (2001). Incentive systems in a real effort experiment. *European Economic Review 45*(2), 187–214.

Van Huyck, J. B., R. C. Battalio, and R. O. Beil (1990). Tacit coordination games, strategic uncertainty, and coordination failure. *The American Economic Review 80*(1), 234–248.

Wilson, T. D., D. A. Reinhard, E. C. Westgate, D. T. Gilbert, N. Ellerbeck, C. Hahn, C. L. Brown, and A. Shaked (2014). Just think: The challenges of the disengaged mind. *Science 345*(6192), 75–77.

7. Experiment 1 Experimental Instructions

Items in italics were not on the subjects screen, but were read aloud.

Costless - Instructions were presented on subjects screens and read out loud by the experimenter.

We will now begin part 2 of the experiment. Please follow along carefully on your screen as we go through the instructions.

In this second part of the experiment, you will participate in 10 rounds of a different decision making task. You will interact in a group for all 10 rounds. Your group consists of 4 people, including you, and the members of your group will remain the same for all 10 rounds of this experiment.

In each round you will be given base earnings of \$0.60. You will then be able to complete instances of a task to generate additional earnings. The task will involve counting the number of 0's in a matrix of 1's and 0's.

Please press continue.

As explained, in each round you will have 2 minutes to complete as many instances of this counting task as you wish. Your payoff will be determined by the number of tasks you complete and the number completed by others in your group.

For each task correctly completed, you will earn a token. You and the other three group members will each receive \$0.20 per token up to the minimum number of tokens earned by any member of your group. For example, if you earned 7 tokens for your group, and the other three group members earned 5, 6, and 7 tokens for the group, the minimum earnings for the group is 5. This means that you and all group members would receive 5x\$0.20 = \$1.00 from the group account. If instead, one group member had earned 0 tokens on behalf of the group, then all members would receive \$0 from the group account, since the minimum earned is 0. Alternatively, if all group members had matched your earnings of 7 tokens, the minimum of the group would be 7, and all members would receive 7x\$0.20 = \$1.40 from these earnings for the group account.

At the bottom of your screen you can see that "Total payment in each round is equal to: \$0.60 + group account payoff.

Please note that these examples used to explain the payoffs are only examples and are not intended to be suggestive of what decisions you should make.

Please press continue for more information on total payment for a round and a few examples.

What you can see on your screen now is a table that shows you the payoff consequences of any choice you or your group members might make up to 7 tokens. You can of course choose to earn more than 7 tokens for the group account and if you do the pattern will remain the same for higher amounts where each additional token that raises the minimum in the group account generates an additional \$0.20.

On the left-side of the table, you will see the sample possible token levels for what you could earn for the group account, or 0-7.

On the top of the table we are giving the minimum earnings for the group account by someone in the group, which also ranges from 0 to 7, though as is the case with you, your other group members could choose to earn more than 7 tokens for the group account.

Example 1:

To read the table, you can just look up a row corresponding to a potential number of tokens you might complete, 7 for example, and then look up the entry in the column for the potential number of tokens you think would be the minimum earned by one of the members of your group. If the minimum of anyone in your group is 7, then your total payoff would be $2.00. This is because your base payoff of $0.60 for the round is added to the additional payoff from the group account that we calculated before, 7*$0.20 = 1.40, to yield a total payoff of $2.00.

Example 2:

As another example, assume you earned 7 for the group account but the minimum earned of the other group members was 6. If you look one column to the left along the row of 7, you will see that your payoff would be $1.80, which is your base earnings of $0.60 plus 6x$0.20 = $1.20 from the group account.

If instead you had earned 6 for the group account, and the other group members earned 6 for the group account, you can look one row up and see that you would still make $1.80, as your total payment would be $0.60 + 6x$0.20 = $1.80.

Example 3:

Finally, assume you earned 1 for the group, while the others all earned more. In this case, your payoff would be $0.80 which is your base payoff of $0.60 and then 1x$0.20 from the group account as the minimum earned for that account was your choice of 1.

Again, the examples we have provided to explain the payoffs should not be taken as recommended actions. We chose these examples simply to help you understand the payoffs from different choices by you and your group members.

Please press continue to learn about the interface you will be using for the duration of the experiment.

At the bottom of the screen you will see a container representing the group account. As you earn tokens through the counting task, those tokens will be deposited into the group account generating earnings as described before.

At the top left of the screen you will see the area where your counting task is. This task requires you to count the number of zeros in the matrix. Once you have completed counting the number of zeros, input this answer into the blue box and press submit. If your answer is incorrect, you will receive a pop-up notice to redo the task. If your answer is correct, once you press submit, a token will be earned. The containers will update automatically with the number earned. Please input the number of zeros you see in this matrix and submit your work to see how this works. <wait until everyone types a line and submits the token – the answer is15 zeros>

After the task is completed successfully and a token deposited, you will be asked to press "next" for the next task. Please do so now.

As this is an example screen, we aren't changing the example, but in a production round you would see a new matrix.

You will have 2 minutes to complete as many of these tasks as you like. Each correctly completed counting zeros task will generate one token that will generate earnings into the group account. On the right side of the screen you will see a clock currently set at 120. This displays the time in seconds and will count down during the paid rounds.

We have kept the payoff table fragment on the upper right portion of the screen for your reference, but note again that you can complete as many instances as you like including going above 7.

At the top of the screen, we have a reminder of how total payoffs for the round are calculated.

Please press continue to see a sample of a results screen for a round.

You will now see an example results screen. Assume for this example you had earned 5 tokens for the group account. Assume also that the other members of your group earned at least 5 tokens for the group account so the minimum earned is 5.

On your screen, you will find a reminder of tokens earned for the group account, which for this example was 5, and below this the minimum earnings to the group account by your group members, which is also 5. Your payoff from the group account is $1.00, which is $.20 times the minimum of 5.

Your total payment for the round would be equal to your fixed pay of $0.60 + group payoff of $1.00, for a total payoff of $1.60.

Please note that this was an example to demonstrate round results and is not suggestive of how much you will actually make in each round.

Please press continue for some final information before we begin the paid portion of this experiment.

Final Information:

You will participate in 10 rounds.

After all decisions have been made in a round, you will always be presented with the results of that round. The computer stores your payment after each round and the total payment from all rounds will be paid to you at the end of the experiment.

Your group will remain the same for the duration of the experiment.

Please press continue.

We are about to begin the paid rounds. Before each round, you will see a screen like this to initiate the next round. You will be directed to the next screen where you will work on the task by clicking on the group account container.

Please click on the group container to begin.

If you have a decision to make, please make it and click on the appropriate buttons so that the experiment will progress. Everything is now self-paced.

8. Experiment 2

Experimenter comments in italics. The treatment differences only appeared in the different wage rates, which we denote in parentheses. Bolded items were also bolded on subjects instructions.

Thank you for participating in today's experiment. I will read through a script to explain to you the nature of today's experiment as well as how to navigate the computer interface with which you will be working. I will be using this script to make sure that all sessions of this experiment receive the same information. Any data collected in today's experiment will be used only for research purposes.

This is an experiment in decision-making. In addition to a $5 participation fee for showing up on time, you will have the opportunity to earn more money through your decisions and the decisions of others in a manner that we will explain soon. You will be paid privately, by cash, at the conclusion of the experiment. All monetary amounts you will see in this experiment will be denominated in dollars and cents. If you have any questions during the experiment, please raise your hand and wait for an experimenter to come to you. Please do not talk, exclaim, or try to communicate with other participants during the experiment. Participants intentionally violating the rules may be asked to leave the experiment with only your show-up payment. At this time, I want to ask you to silence your cell phones and put them away.

Description of Experiment

There will be two separate parts of today's experiment involving completely separate and unrelated decision tasks. You will go through each part separately meaning that after we have gone through instructions for Part 1, you will make decisions in this part. Once Part 1 is complete, we will go through instructions for Part 2 and you will then be able to make decisions in Part 2. Your earnings from each part of the experiment will be added together to generate your total for the experiment.

Part I

This part will be broken into two phases. The first phase will involve one round of your engaging in a task after which we will give you additional instructions to engage in 3 more similar rounds. Please press the continue button on your screen now and it will take you to an example screen of what you will be doing in the first phase.

Phase I

During the paid round, you will see a screen with 48 bars, though on this practice screen there are only 5. We will call these bars sliders as on each bar there is a marker you can slide along the bar with your mouse. To move the marker you can click on the slider or you can click on the marker and drag it along. The markers on each slider are located at the 0 position. You can see the position for each slider to the right of the slider. Your task will be to set the markers on as many sliders as you choose to a position of 50. You will have 3 minutes to do so.

For each slider you set to 50, you will be paid $0.01 ($0.04). This means that if you are able to complete 15 sliders then you will earn $0.15 ($0.60). If you complete more your earnings will be higher and if you do fewer your earnings will be lower. At the end of the round, you will be shown how many sliders you correctly aligned as well as your earnings for the round.

At the top right of your screen, you will see a countdown clock that will start at 180 seconds and countdown until the end of the round. At the middle of the top of your screen you will see how many sliders you have correctly aligned and your payment information.

For this practice screen, please align all five sliders to 50 so that you see how to do it. Once you have completed all five, a continue button will show up at the bottom. Press this and you will be taken to a screen summarizing the information for the next round.

Once everyone has pressed the continue button from that screen, we will begin the first round for Phase 1 of Part 1 of the experiment which will last for 3 minutes.

Are there any questions?

Phase II

The first phase has ended and you can see your earnings from this phase. Please press OK and we will go through the instructions for the second phase of this first part of the experiment.

The second phase consists of three, 3-minute rounds. In this second phase, you will again be presented with a screen of sliders to align and again you will earn $0.01 ($0.04) for each one correctly aligned. We will call this screen 1. You will also have the opportunity to engage in some alternate activities. These alternate activities will be available on screen 2. Unlike your payment from screen 1, your screen 2 payment is based on the total time spent on screen 2, not on how many activities you complete there. In each of the rounds, you can spend all of your time on screen 1 and none on screen 2, all of your time on screen 2 and none on screen 1 or a mix of time spent on both. The screen you see in front of you explains how your screen 2 payment will work.

Each round in this phase will begin on screen 1 where you can align sliders for $0.01 ($0.04). You will be able to align sliders for as much of a 3 minute round as you wish and you can switch to screen 2 at any point. If you want to switch to screen 2, there will be a button allowing you to do so. If you switch immediately and spend all 3 minutes on screen 2 you will earn $1.19 for that round. The total payment for screen 2 will fall the longer you spend on screen 1 aligning sliders. Specifically you will lose an amount from the $1.19 payment for every second you spend on screen 1 aligning sliders.

In order for you to more easily see how these costs work, this graph shows what the cost will be for each 5-second increment you spent on screen 1 aligning sliders.

Let's go through some examples. On your screen, there is a slider bar next to the "Time Spent" label that you can slide along the graph to see what the additional cost of each 5-second increment will be. This slider bar represents how much time you could spend on screen 1. Initially the cost of spending 5 more seconds aligning sliders is low but it increases as you spend more time on screen 1 aligning sliders. You can see that the first 5 seconds costs 0 cents. As you increase the time spent on screen 1 (represented by moving the slider), you will see that the cost per 5-seconds spent on screen 1 rises to 0.6 cents at 25 seconds, 1.4 cents at 50 seconds, 2.4 at 75 seconds and the cost for 5 seconds aligning sliders would rise up to 6 cents at 150 seconds. Remember, this is not the total cost of spending time aligning sliders, but rather the cost of spending an additional 5 seconds aligning sliders. The total cost represents the sum of the costs up to that point in time. As you increase the time you spend on screen 1 aligning the sliders, notice that the total cost you must pay for spending time on screen 1 is also updated. Please use the slider bar to explore this cost function so that you have an understanding of how it works and click the continue button when you are ready to move on. As we will show you on the next screen, this information will be available on your decision screen. Please raise your hand if you have any questions about how your payment works.

You now see an example version of screen 1, which again will contain 48 sliders during the paid rounds. Again, for the example screen where there are only 5. There are a few additions to this screen over the version for the previous phase. First, you will note the button to "Go to Screen 2". During a production round, you will be able to click on this button to switch to screen 2 but once you choose to go to screen 2, you will not be able to come back to screen 1 in that round. For this practice screen, you will be able to switch back and forth freely. At the top, you will also see information about your potential screen 2 earnings along with the information about your current screen 1 earnings.

In the center, you will see a few useful bits of information. First, we have provided an estimate of what your earnings will be for 5 seconds of aligning sliders based upon your speed of slider alignment from phase 1. This is just an estimate and if you are faster or slower than your average speed from that round, your earnings may be higher or lower. Below this you can see the cost of aligning sliders for the next 5 seconds. For this practice screen we have included a slider bar that you can use to move along to see the same information you did on the prior screen about what the cost would be for any five second increment. Again the first 5 seconds costs you nothing, at 25 seconds the cost rises to 0.6 cents for a five second increment, at 50 seconds to 1.4 cents, 2.4 cents at 75 seconds and 6 cents at 150 seconds. At the top you will see your potential screen 2 earnings decline the longer you spend on screen 1 which on this screen, can be seen by moving the slider along that time bar.

Just to confirm how earnings work with a specific example, if you spend 75 seconds aligning sliders and then switch to screen 2 the total cost from your screen 2 earnings would be 15.9 cents which is subtracted from your 119 cents meaning you would earn 103 cents from time spent on screen 2. If you aligned 15 sliders during those 75 seconds you would earn 15 cents (60 cents) from slider alignment meaning that your total earnings would be 118 cents (163 cents).

Remember, during a paid round, you can switch to screen 2 when you decide you no longer want to align sliders. To see the activities available on screen 2, please press the button to go to screen 2 now.

You should now see a sample of screen 2. While on this screen, you can choose to play Tic-Tac-Toe (TTT) against a computer algorithm on the left side of your screen or you can do a word jumble on the right side. In TTT, you will be X's and the computer will be O's. You and the computer will take turns placing your X's and O's in the empty cells on a 3x3 table. Your goal is to get three X's in a row either across or diagonal. If you do before the computer gets three O's in a row, you win. If the computer get's 3 O's in a row, they win. For the word jumble, you see a matrix of letters. Inside that matrix are ten hidden words going across, down, backwards or diagonal. The category of the word jumble is given at the top of your screen. For instance, the first category is colors. When you see a word you can type it into the text box at the lower right and click on the submit button. Once you have found all of the words in a given word jumble, a new one will appear with a different category.

These activities will be available for you to engage in as you wish. Remember, there are no earnings for completing these activities because compensation for screen 2 is calculated by the time spent on screen 2. Once you switch to this page, you will stay on it until the round ends.

Summary: During a round, you will have 3 minutes to align as many sliders as you choose receiving $0.01 ($0.04) for each one aligned. For each second you spend aligning sliders, you will pay a cost in the form of a reduction of your screen 2 earnings as we explained before. This cost, represented as the additional cost for five more seconds spent aligning sliders, is always available to you at the top of your screen. At any time if you do not wish to continue aligning sliders, you will be able to switch to screen 2 and preserve your remaining screen 2 earnings while engaging in either TTT or the word jumble for the remainder of the time. Once you have switched over to screen 2 in a round, you will not be able to switch back to aligning sliders, though you will restart on the slider alignment screen in the next round. This second phase will involve three, 3-minute rounds.

Are there any questions about this phase?

You now see a screen reminding you of the details of this round. We have also provided you with information regarding how much you should expect to earn per five seconds of aligning sliders if you continue aligning them at the same speed as in Phase I. You may of course align them faster or slower in subsequent rounds and so this estimate may not reflect your actual earnings but it is an estimate based on your earlier speed of slider alignment. Please press continue when you are ready to move on.

9. Experiment 3 Experiment Instructions

9.0.1. Verbal Script

Items in Bold represent the instructions presented on a single screen. The <> represents the beginning and end of different instructions given in the different treatments. Items in italics were not on the subjects screen, but were read aloud.

Welcome.

This experiment is a study of group and individual exchange behavior. The instructions are simple, and if you follow them carefully and make good investment decisions, you may earn a considerable amount of money, which will be paid to you at the end of the experiment.

Please wait while I start the system.

Please follow along carefully on your screen as we go through the instructions.

<Trivial Real Effort>

General Information:

You are one person in a group of 4.

Each of you will complete a task to earn a set of 10 tokens per round. The task is similar to data entry. We will provide you with printouts of the data and we would like you to enter it into the computer.

All 10 tokens that you earn will be invested to turn them into income.

There are two ways you can earn money by investing your tokens.

Please press continue to find out more about these two ways of investing your tokens.

</end Trivial Real Effort>

<Useful Real Effort>

General Information:

You are one person in a group of 4.

Each of you will work to earn a set of 10 tokens per round. The task you will be engaged in is entering data on mutual fund investments. We will provide you printouts of the data and we would like you to enter it into the computer so that we can create an electronic data file with this information. This data is required by another researcher who will be using the data entered here for a research project on financial markets. Since this data will be used in a research project it is important that it be entered in accurately and completely.

All 10 tokens that you earn will be invested to turn them into income.

There are two ways you can earn money by investing your tokens.

Please press continue to find out more about these two ways of investing your tokens.

</end Useful Real Effort>

<Stylized>

General Information:

You are one person in a group of 4.

Each of you will be given a set of 10 tokens per round.

All 10 tokens that you receive will be invested to turn them into income.

There are two ways you can earn money by investing your tokens.

Please press continue to find out more about these two ways of investing your tokens.

</end Stylized>

The first way you have to invest is the Individual account, which you can see described on this screen.

The Individual Account:

For every token you invest in the individual account, you will receive $0.20.

For example, if you invest 4 tokens in the individual account you will earn $0.80 (4x$0.20)

If you invest all 10 tokens in the individual account, you earn $2.00.

Please press continue to learn about the second option you have for investing your tokens.

The second way you have to invest is the group account.

The Group Account:

The return on your investment in the group account is not so easily determined.

Your earnings depend upon the total investment in the group account (your invested tokens plus all of the other tokens invested by the other people in your group.)

Please click on the Continue button to find out more about the group account.

The Rules of the Group Account:

You and 3 other people are members of a group. In your group, you and all other members have (earned) 10 tokens.

Every token invested by a group member in the group account generates $0.40 to the group and all group earnings are split equally.

This means that you and the other group members earn $0.10 for every token invested into the group account. It does not matter who invests the tokens in the group account. All earnings are divided equally between all members.

The more the group invests, the more all group members earn from the group account.

Please press continue to see an example.

Group Account Example:

If you put no tokens in the group account, but the other members invested enough to earn $2.00, you would earn $0.50 from the group account.

If the others invest no tokens in the group account, but you put in enough for the group to earn $2.00, you still earn $0.50 from the group account.

Remember – how much you and the other group members earn depends on how much all of you together put into the group account.

You have been given a handout with a payoff table outlining how tokens invested in the group account translate to earnings. Please turn to this handout now as we go through example 1.

Example 1:

Suppose that the members of your group together invest a total of 24 tokens in the group account. From the table, individual and group payment corresponds to the row where 24 tokens are contributed. You will see from the table that the group earns $9.60 and each group member earns their equal share of this amount, $2.40.

Notice that you would receive $2.40 from the group account whether or not you invested any of your tokens into the group account.

Now that we have discussed both investment options for your 10 tokens, please press continue to learn about the interface you will be using for the duration of the experiment.

<UE and TE>

At the bottom of the screen you will see two containers, which represent the two accounts, individual and group, that you can invest your tokens into. To select an account to invest in, you must click on the container. To see how this works, please select the individual account by clicking on the yellow container. You will see that the group account container is now grey which indicates your only active account is the individual account. You will see a reminder of the $0.20 payoff for each token invested in the individual account to the left of the individual account container.

At the top of the screen you will see the area where you will input your data. You have been given a sheet of paper with the data to be input. For now, please see the handout with "Practice Work" written at the top. In the appropriate sections, please carefully type the data from the sheet into the boxes provided for the specified categories.

<UE only>

Please remember that the data you are inputting is required by another researcher who will be using the data you input for a research project on financial markets. Since this data will be used in a research project it is VERY important that it be entered in accurately and completely.

</end UE only>

You only need to type in the first line of data. Once you have finished this, submit your work by pressing the submit button. This will generate a token, which will be automatically invested in the active account, which is currently the individual account. The containers will update automatically with the number invested. Please input the data and submit your work to see how this works.

Please click on the group account container now. You will see that it has highlighted orange and the individual account is now gray. This indicates that the group account is now active and any tokens will be invested in this account. You will see a reminder of the \$0.40 payoff to the group for each token invested in the group account to the right of the group account container. Please type in the second line of data and then click on the submit button.

You will work to earn 10 total tokens in a round. Each data line will generate one token, which you will choose how to invest. You will see how many tokens are remaining to be earned and invested which is laboratory eled "investments left to make."

Please press continue to see a sample of a results screen for a round.

</end UE and TE>

<SE>

Tokens will arrive at certain intervals. How these tokens are invested depends upon which account you currently have activated. At the bottom of the screen you will see two containers, which represent the two accounts, individual and group, that you can invest your tokens into. To activate an account to invest in, you must click on the container. To see how this works, please select the individual account by clicking on the yellow container. You will see that the group account container is now grey which indicates your only active account is the individual account. You will see a reminder of the \$0.20 payoff for each token invested in the individual account to the left of the individual account container.

At the top of the screen you will see a tic-tac-toe board on the top right and on the top left, at some point, you should see or have seen a green token appear. This will be automatically invested in the active account, which is currently the individual account. You can change the active account at anytime by clicking on the other container. Tokens will appear at random intervals though you will be warned 3 seconds before they are automatically invested in case you want to redirect it to a different account. The containers will update automatically with the number or tokens invested in each.

Please click on the group account container now. You will see that it has highlighted orange and the individual account is now gray. This indicates that the group account is now active and any tokens will be invested in this account. You will see a reminder of the \$0.40 payoff to the group for each token invested in the group account to the right of the group account container.

You will be given 10 total tokens in a round, which you will choose how to invest by clicking on the respective containers. You will see how many remaining tokens are coming in red above the containers "investments left to make."

Tic-tac-toe is provided for you to play if you wish to while you are waiting on tokens. You will be playing against a computerized opponent and there is no payment for winning or losing a tic-tac-toe game.

Please press continue to see a sample of a results screen for a round.

</end SE >

You will now see a sample results screen. Recall that you will have 10 tokens in a round. Assume for this example you had invested 5 of these tokens to the individual account and 5 to the group account. Assume also that the other members of your group had invested a total of 16 tokens to the group account, so the total amount of tokens in the group account is 21.

At the top you will find your individual earnings, which for 5 tokens invested would be equal to $1.00.

Below this you will find a reminder of your contribution to the group account, which for this example was 5, and below this the sum of all contribution by all members (including you) to the group account. In this example, the total number of tokens contributed by all members of the group is 21, meaning the group earned a total of $8.40. Your share of the group earnings would be $2.10.

Your total earnings for the round would be equal to your individual earnings of $1.00 plus your share of the group earnings, $2.10, for total earnings of $3.10.

Please note that this was an example to demonstrate round results and is not suggestive of how much you will actually make in each round.

Please press continue for some final information before we begin the paid portion of this experiment.

Final Information:

You will have 10 rounds in which you will make investment decision about your 10 tokens. The profits you make are stored by the computer after each round and the total profits from all rounds will be paid to you at the end of the experiment.

After you and everyone else in your group has had a chance to invest all 10 tokens in a round, you will be shown a results screen which will inform you of how many tokens were invested in the group account and how many you have put into your individual account. You will also been shown your own earnings from the group account and from the individual account, and your total earnings for the round.

You group will remain the same for the duration of the experiment.

We are about to begin the experiment. If you have a decision to make, please make it and click on the appropriate buttons so that the experiment will progress. Everything is now self-paced. Please press continue.

We are about to begin the paid rounds. Before each round, you will see a screen like this to initiate your investment account. Once you select the initial investment account (and remember, this can be changed at anytime) you will be directed to the next screen where you will have 10 tokens to invest.

Please make your selection and begin.

9.0.2. Subject Handouts

Tokens:

- Tokens will appear at randomly timed intervals.
- Once the token appears, it will be automatically invested into your selected account after 3 seconds.
- You will be given 10 total tokens in each round.
- There will be 10 total rounds.

Tips on Investing:

- Tokens will arrive at random intervals and will be invested into the account you currently have chosen to be active.
- You can change your chosen account at any time by selecting the alternative choice.
- You can choose to divide the tokens between accounts in any way you want. That is, you can:
  - put all of them into the group account
  - put all of them into the individual account
  - put some of them into the group account and some of them into the individual account.

Earnings:

- The individual account pays $0.20 per token you invest.
- The group account pays $0.40 per token invested by each member of the group. These earnings will then be split equally between all 4 members of the group meaning you generate $0.10 to yourself and $0.10 to every other member of the group.
- You do not have to invest in the group account to receive earnings from the group account.

Stylized Effort Cheat Sheet – handed out along with payoff table

Practice Work:

	Ticker	Code	1-yr return	5-yr return	10-yr return
1.	MKEGH	AD	87.86	70.71	77.75
2.	UULIF	AE	-95.16	-3.26	-89.14

Work:
- Please work to be as accurate in you data entry as possible
- If it says N/A – please input 999
- Only input Ticker, code, 1 Yr Return, 5 Yr Return, and 10 Yr Return (ignore the categories NAV, Total Net Assets and Since Inception)
- You will work to earn 10 tokens in each round.
- There will be 10 total rounds.

Tips on Investing:

- When you press the "submit work" button after each data entry, you will automatically generate a token which is invested in the account you currently have chosen to be active.
- You can change your chosen account at any time by selecting the alternative choice.
- You can choose to divide the tokens between accounts in any way you want. That is, you can:
  - put all of them into the group account
  - put all of them into the individual account
  - put some of them into the group account and some of them into the individual account.

Earnings:
- The individual account pays $0.20 per token you invest.
- The group account pays $0.40 per token invested by each member of the group. These earnings will then be split equally between all 4 members of the group meaning you generate $0.10 to yourself and $0.10 to every other member of the group.
- You do not have to invest in the group account to receive earnings from the group account.

Trivial Effort Cheat Sheet – handed out along with payoff table

Practice Work:

Ticker	NAV as of 10/5/2015	Total Net Assets	Load Adjusted Returns			
			1 Yr Return	5 Yr Return	10 Yr Return	Since Inception
ADVWX	11.47	$14,700,000	-7.28%	9.11%	N/A	9.38%
ABAGX	N/A	N/A	0.90%	6.26%	2.35%	2.59%

Work:

- Please work to be as accurate in you data entry as possible. Your work will help to create a database to be used by researchers and so accuracy is highly important.
- If it says N/A – please input 999
- Only input Ticker, code, 1 Yr Return, 5 Yr Return, and 10 Yr Return (ignore the categories NAV, Total Net Assets and Since Inception)
- You will work to earn 10 tokens in each round.
- There will be 10 total rounds.

Tips on Investing:

- When you press the submit work button after each data entry, you will automatically generate a token which is invested in the account you currently have chosen to be active.
- You can change your chosen account at any time by selecting the alternative choice.
- You can choose to divide the tokens between accounts in any way you want. That is, you can:
  - put all of them into the group account
  - put all of them into the individual account
  - put some of them into the group account and some of them into the individual account.

Earnings:

- The individual account pays $0.20 per token you invest.
- The group account pays $0.40 per token invested by each member of the group. These earnings will then be split equally between all 4 members of the group meaning you generate $0.10 to yourself and $0.10 to every other member of the group.
- You do not have to invest in the group account to receive earnings from the group account.

Useful Effort Cheat Sheet – handed out along with payoff table

Group Account Payoff Table

Tokens invest in the GROUP account by ALL group members	Total Money Earned by the GROUP	How much money YOU earn
0	$0	$0
1	$0.40	$0.10
2	$0.80	$0.20
3	$1.20	$0.30
4	$1.60	$0.40
5	$2.00	$0.50
6	$2.40	$0.60
7	$2.80	$0.70
8	$3.20	$0.80
9	$3.60	$0.90
10	$4.00	$1.00
11	$4.40	$1.10
12	$4.80	$1.20
13	$5.20	$1.30
14	$5.60	$1.40
15	$6.00	$1.50
16	$6.40	$1.60
17	$6.80	$1.70
18	$7.20	$1.80
19	$7.60	$1.90
20	$8.00	$2.00
21	$8.40	$2.10
22	$8.80	$2.20
23	$9.20	$2.30
24	$9.60	$2.40
25	$10.00	$2.50
26	$10.40	$2.60
27	$10.80	$2.70
28	$11.20	$2.80
29	$11.60	$2.90
30	$12.00	$3.00
31	$12.40	$3.10
32	$12.80	$3.20
33	$13.20	$3.30
34	$13.60	$3.40
35	$14.00	$3.50
36	$14.40	$3.60
37	$14.80	$3.70
38	$15.20	$3.80
39	$15.60	$3.90
40	$16.00	$4.00

Payoff Table

10. Appendix

	(1) BRL1	(2) Tobit1	(3) BRL2	(4) Tobit2	(5) BRL3	(6) Tobit3
TE	0.465	0.690	-0.175	-0.334	-0.908	
	(0.932)	(1.035)	(0.802)	(1.449)	(1.350)	(1.874)
UE	-0.268	-0.427	-0.712	-1.050	-0.690	-1.124
	(1.058)	(1.192)	(0.800)	(1.701)	(1.649)	(2.040)
Period			-0.423***	-0.686***		
			(0.111)	(0.170)		
Period∗TE			0.116	0.190		
			(0.155)	(0.199)		
Period∗UE			0.0807	0.120		
			(0.148)	(0.213)		
Group_{t-1}					0.178***	0.324***
					(0.064)	(0.076)
Group_{t-1}∗TE					0.063	0.064
					(0.069)	(0.097)
Group_{t-1}∗Useful					0.038	0.080
					(0.103)	(0.121)
Constant	5.454***	5.733***	7.781***	9.461***	2.285*	-0.0914
	(0.674)	(0.808)	(0.354)	(1.241)	(1.284)	(1.533)
Obs (Clusters)	880 (22)	880 (88)	880 (22)	880 (88)	792 (22)	792 (88)

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

BRLX: Biased-Reduced Linearization clustering at the group level to correct for the small number of clusters.

TobitX: Tobit with bounds [0,10] with standard errors are clustered at the subject level..