The Pennsylvania State University

The Graduate School

College of the Liberal Arts

MULTI-MODAL EFFORT ESTIMATION

A Thesis In
Psychology
by
Cory Adam Potts

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

May 2016

The thesis of Cory Adam Potts was reviewed and approved* by the following:

David A. Rosenbaum
Distinguished Professor of Psychology
Thesis Adviser

Richard A. Carlson
Professor of Psychology
Associate Head of the Department of Psychology

Karen Gasper
Associate Professor of Psychology

Melvin Mark
Professor of Psychology
Department Head

*Signatures are on file in the Graduate School

ABSTRACT

Cognitive and physical tasks appear to have little in common, yet we are able to make decisions such as which of two tasks is easier or which to do first. Our ability to do so suggests some common variable(s) used to compare different kinds of cost. Recently, we discovered a phenomenon that bore on this topic. We asked participants to choose to either carry a bucket a short walk away or carry a bucket a long walk away to equidistant end points. Surprisingly, participants consistently carried the closer bucket a longer distance, thereby expending unnecessary physical effort. We ascribed this tendency to *pre-crastination*, which we defined as the tendency to hastily offload tasks from working memory, even at the expense of added physical effort. In Experiments 1 and 2 of the present thesis, I tested whether participants were willing to reach long distances and heft heavy buckets to show pre-crastination. Despite the need to lean and reach far, participants continued to favor the closer bucket. However, the probability of picking up the closer bucket fell to chance when it contained a heavier weight load (7 lbs.), while the farther bucket contained a lighter weight load (3.5 lbs.). In Experiments 3-6, I more directly tested the relation between cognitive and physical effort by asking participants to choose to perform either a mainly physical task (reaching for a bucket with different physical weight loads and carrying it some distance) or a mainly cognitive task (counting to a target number). I did so with two aims. The first was to build a metric to relate cognitive and physical costs. The second was to test the hypothesis that measured time is the substrate that relates cognitive and physical costs. The results were consistent with the hypothesis that participants used time as an index of effort. However, there was evidence to suggest that they rely on time less as task demands increase. I discuss potential alternatives to the use of time to relate cognitive and physical effort in the General Discussion section, with a specific focus on the idea that subjective effort is the amount of monitoring one must do throughout a task.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to offer a special thank you to my committee, David Rosenbaum, Rich Carlson and Karen Gasper, for their support throughout this process. I would also like to thank Iman Feghhi, John M. Huhn III, and Teenu Sarjaveen for helpful comments. Finally, I would like to thank my truly excellent team of research assistants (listed in alphabetical order), Brielle Bailey, Emily Doñe, Alexa Echevarria, Meghan O'Brien, Mauli Patel, Carly Rosenthal, Dan Sheng, and Genevieve Stafford.

CHAPTER 1. INTRODUCTION

Which is easier: memorizing seven digits or lifting seventy pounds of weight? Although these tasks share little in common, you were likely able to answer this question rather easily. You probably chose the former activity over the latter. This thought exercise, as well as the daily decisions that inspired it, raise an interesting question: How we are able to weigh the effort involved with cognitive tasks and physical tasks to make decisions such as which is easier or which should be done first. If comparing cognitive and physical tasks were the "apples-and-oranges" decision that it intuitively seems to be, then it would be impossible to make decisions of this kind. Our ability to do so suggests that these two domains share a common unit of measure.

Recently, we (Rosenbaum, Gong, & Potts, 2014) discovered an interesting phenomenon in the planning of physical actions that bore on this "apples-and-oranges problem." We asked 257 university-student participants to walk down an alley without stopping, to pick up either a left or a right bucket, whichever seemed easier, and to carry that bucket to a corresponding far table. In a typical critical trial, participants chose to either carry a closer bucket a longer distance or carry a farther bucket a shorter distance. We expected participants to choose to carry the farther bucket a shorter distance, thereby minimizing physical effort. Surprisingly, the vast majority of participants did the opposite. They chose to carry the closer bucket a longer distance, even at the expense of added physical effort. This finding remained true even when both buckets contained a relatively heavy load (7 pounds of pennies). Inspired by the consistent self-report given by participants that they chose the closer bucket to "get it done as quickly as possible," we ascribed the close-bucket preference to "pre-crastination," a word we invented to mean the hastening of sub-goal completion at the expense of added physical effort. Contrary to the phenomenon of procrastination, in which goals are typically delayed, our participants elected to complete sub-goals sooner than necessary.

We considered two interpretations for pre-crastination, both of which highlight the tradeoff between different kinds of costs. The first interpretation involves working memory demands. It is possible that participants chose the closer bucket to more quickly eliminate a sub-goal from working memory, namely, selecting which of the two buckets to pick up. This interpretation is similar to the idea of *offloading* in the field of embodied cognition, in which features of the environment are used to reduce cognitive load (Wilson, 2002). For example, while playing the game of Tetris, players tend to use physical rotations to explore shapes and plan strategies rather than relying on mental rotations (Kirsh & Maglio, 1994). In this case, participants may have been able to offload mental effort more quickly if they picked up the closer bucket, though this resulted in the unnecessary expense of physical effort.

The second interpretation for pre-crastination involves sign-tracking, a construct rooted in Pavlovian theory. Sign-tracking, or autoshaping, involves an arbitrary association between a stimulus and the onset of a reward. For example, rats will gnaw or lick a lever associated with the delivery of a food reward (Davey & Cleland, 1982). Some researchers have suggested sign-tracking as the mechanism that causes drug addicts to respond physiologically to tools associated with the delivery of drugs, such as glass bottles or hypodermic needles (Tomie, Grimes, & Pohorecky, 2008). In the case of our experiments, participants may have associated the closer bucket with an increase in positive affect due to the quick completion of a sub-goal. It may have felt good to advance more quickly toward the ultimate goal of dropping one of the buckets off at its corresponding far table. This interpretation is consistent with previous research suggesting that people are willing to work harder to complete activities associated with positive affect (Custers & Aarts, 2005).

Though distinguishing between these two theoretical accounts is of interest, two larger conclusions drawn from the pre-crastination series are more central to this master’s thesis. The first conclusion is that, whether cognitive, affective, or physical, our participants were able to choose between different kinds of costs. They were able to resolve, or perhaps did not even notice, an “apples-and-oranges” problem. For the purposes of this proposal, the most relevant “apples” and “oranges” will be cognitive and physical costs. The second conclusion is that this series suggested the use of observable physical actions to study cognitive cost. To the best of my knowledge, this is the first time that such a comparison was made.

In the master's thesis to follow, I had two primary goals. The first was to test different kinds of costs within a single modality. In Experiments 1 and 2, I tested the relative cost of reaching some distance versus walking some distance for objects with various physical properties. The second goal was to test different kinds of costs across multiple modalities. In Experiment 3, I compared physical and cognitive effort with the aim of building a metric to relate these two kinds of costs. In Experiments 4-6, I tested whether participants used the difference in measured time between two tasks to relate cognitive and physical costs across increasing task demands.

Comparing Physical Costs

If physical actions are to be used as a proxy for cognitive cost, then it is important to establish a reliable method for measuring physical costs. Some investigators have approached the study of physical cost using self-report ratings. Such ratings have a long history in the study of physical exertion (Robertson & Noble, 1997). A typical experimental paradigm involves varying the intensity and complexity of the physical tasks that participants perform, and then asking for corresponding ratings of exertion. The most famous and widely used measure of exertion is the Borg Rating of Perceived Exertion (RPE) scale (Borg, 1962). RPE ratings have been validated using physiological measures of exertion such as oxygen uptake, relative peak oxygen uptake, pulmonary ventilation, heart rate, respiratory rate, and respiratory-exchange ratio (Guidetti et al., 2011).

The success of self-report ratings in the study of physical exertion raises the question of whether similar methods could be used to study the underpinnings of subjective effort. Rosenbaum and Gregory (2002) pursued this possibility by asking participants to perform oscillations about the elbow in the horizontal plane at prescribed frequencies and amplitudes, and then to assign retrospective ratings from 1-5 to the action, where "1" denoted least effort and "5" denoted most effort. The authors expected ratings of effort to increase with greater movement amplitudes and higher frequencies. Said another way, they expected that greater angular displacements and faster rates of movement would correspond to increased feelings of effort. Indeed, the authors found that movements performed at higher frequencies were given higher ratings of perceived effort. However, a surprising relation emerged between movement amplitudes and ratings of perceived effort. Contrary to the authors' expectations, increases in movement amplitudes corresponded to *lower* ratings of effort. Rosenbaum and Gregory suggested that this unexpected relation might have been a product of feedback about movement accuracy given to participants at the end of each trial. Participants may have reasoned that when a trial ended in positive feedback, it must have been less effortful, whereas when a trial ended in negative feedback, it must have been more effortful. This unexpected result highlights the importance of psychological factors in the experience of physical effort.

Although self-report ratings have been a successful tool in the study of physical effort, a relevant question is not only how features of actions relate to movement cost, but how different kinds of physical actions relate to one another in terms of cost. The successful use of two-alternative forced choice (2AFC) in the study of action planning (for a review, see Rosenbaum et al., 2012), led Rosenbaum (2008) to test the relative costs of different physical actions, namely reaching some distance versus walking some distance, using a 2AFC design. Rosenbaum asked right-handed participants to face a table with a single bucket placed on its surface. The bucket could have been placed close to the left edge, in the middle, or close to the right edge of the table. Two barstools, one in line with the left edge of the table and the other in line with the right edge of the table, functioned as target destinations for the bucket. The walking distances to the left and the right stools also changed across trials. Participants were instructed to do what seemed easier: walk around the left side of the table, pick up the bucket with their right hand, and carry it to the left stool, or walk around the right side of the table, pick up the bucket with their left hand, and carry it to the right stool. The probability of walking around the right side of the table was found to equal the proportion of *functional distance* to the left stool, where functional distance included: (1) the walking distance (m) to either the left stool or the right stool; (2) the reaching distance (m) from either side of the table to the bucket's handle; (3) a coefficient to represent the cost of reaching relative to walking per unit distance to be multiplied by the reaching distance.

The strongest correlation (r = .97) between the predicted and observed probabilities of walking around the right side of the table was achieved when the left hand reaching cost was set to 12.3 and the right hand reaching cost was set to 10.3. Said another way, reaching one meter with the left hand was equivalent to walking 12.3 meters of distance, and reaching one meter with the right hand was equivalent to walking 10.3 meters of distance. Averaging across the two hands, reaching over a meter of distance was approximately 11.3 times costlier than walking over that same distance, suggesting that lateral reaches are relatively costly as compared to steps forward. The added expense for left hand reaches was sensible, as the majority of participants were right-handed, as indicated by a 7 or higher out of 10, where "1" on the short form of the Edinburgh Handedness Inventory. In this inventory "1" denotes the strongest left-hand preference and "10" denotes the strongest right-hand preference (Oldfield, 1971).

Comparing Cognitive Costs

Drawing upon the success of self-report ratings in the study of physical exertion, some researchers have turned to the use of self-report ratings as an index of cognitive cost (Galy et al., 2012). Specifically, two self-report measures of cognitive effort have been widely employed across the cognitive workload literature. The first, and arguably the most popular, is called the NASA Task Load Index (NASA-TLX; Hart & Staveland, 1988). The NASA-TLX includes six subscales to measure physical demand, mental demand, temporal demand, performance, frustration, and effort. The second commonly used self-report measure is the Subjective Workload Assessment Technique (SWAT; Reid & Nygren, 1988). The SWAT measures three dimensions of task demand, including temporal demand, mental demand, and psychological stress.

Although these self-report measures of cognitive effort are widely used and well validated, there are some issues with this methodology. One issue is that self-report measures are often retrospective, and therefore collapse across potentially interesting non-linear subtleties of effort. This issue has led some researchers to seek appropriate psychophysiological indices for a more sensitive on-line measure of cognitive effort. One measure that is gaining momentum within the cognitive workload literature is pupillary response (Marshall, 2002). As task demands increase, so does the circumference of the pupil. The strong correlation between task difficulty and pupillary dilation has led Marshall and colleagues to develop the Index of Cognitive Activity (ICA), an on-line measure of pupillary activity across a task. A promising feature of the ICA is that, with a sufficiently high sampling rate, one may be able to pick up on subtle non-linear features of effort across the course of a task. Similar efforts have been made to correlate increases in cognitive load with heart rate (Backs, 1995) and specific EEG signatures (e.g., Marshall, 2002).

With these measures of cognitive cost in mind, a relevant question is how different cognitive tasks compare in terms of their relative effort. In pursuit of this question, Robinson and Morsella (2014) tested the cognitive effort associated with attending, assessing, and choosing. For the "attending" task, participants were asked to attend to two shapes, and to report any changes in their form. For the "assessing" task, participants were shown two shapes, and then reported which of the two shapes fit a predetermined criterion (e.g., which shape had more corners). For the "choosing" task, participants were simply asked to choose the shape they liked the most. In each trial, after making their selection, participants rated the effort associated with that trial on an eight-point scale. The authors found that attending was given the highest rating of subjective effort, followed by assessing and then choosing. However, it is important to note that these tasks were not necessarily mutually exclusive; during the "attending" task, participants were, presumably, assessing various features of the relevant shapes. Likewise, during the "assessing" task, participants were very likely attending to relevant features. Across all conditions, responding required choosing. Therefore, although the results are sensible, there was some potential cross-contamination among conditions. One can only speculate that such issues with cross-contamination among cognitive tasks is why so few studies have directly compared different cognitively effortful tasks.

Toward a Common Currency of Effort

A larger conclusion one can draw from reading the previous two sections is that the study of physical effort and the study of cognitive effort are fundamentally similar. Broadly, researchers interested in either form of effort have pursued almost identical methodologies, including self-report ratings of effort and on-line psychophysiological measures. In fact, the close ties between cognitive and physical effort have led some researchers to question whether these two forms of effort are, in fact, separate. Instead, it is possible that effort reflects the depletion of some resource that is common to tasks of different kinds. Gailliot and Baumeister (2007) have suggested this resource is glucose. In support of this hypothesis, these authors have studied what they call *ego depletion* (Baumeister et al., 1998). Ego depletion is the name given to the finding that people tend to perform poorly on self-control tasks after having just exercised self-control. Said another way, self-control is a limited resource that can become depleted when used. For example, the authors found that people who fought the temptation to snack on chocolate chip cookies gave up more quickly on a subsequent impossible-to-solve puzzle than peers who resisted radishes. Interestingly, providing participants with a boost of blood glucose appears to be one way to fight the effects of ego depletion. Supporting the relation between blood glucose and ego depletion, Gailliot et al. (2007) found that participants who carefully attended to specified features of a video subsequently had lower glucose levels than participants who viewed the video as they normally would. Likewise, lower glucose levels predicted poor performance on a Stroop task to follow. However, if participants consumed a sugary drink prior to the Stroop task, the deleterious effects of sustained attention diminished.

However, some authors have questioned the claim that blood glucose is, in fact, the currency of effort (Kurzban et al., 2013). Instead, these authors suggest that a phenomenological account of effort is preferable to a hard-and-fast physiological resource. Indeed, there are findings that challenge a strict resource depletion account of effort. One is that ego depletion can be attenuated, to some extent, without administering glucose. For example, inducing positive affect between self-control tasks can boost performance (Tice et al., 2007). If depleting levels of glucose were the sole source of effort, one would expect to see only impoverishment, rather than improvements, in performance. However, it is conceivable that positive affect somehow influences levels of blood glucose.

In their model of effort, Kurzban and colleagues (2013) suggested that the functional purpose of effort is to divert attention away from a current activity and direct it instead to other potentially beneficial activities. According to these authors, felt effort is the asymmetry between the benefits of continuing the task at hand and the potential benefits of switching to a new task. If, for example, there were an exciting concert happening right now, the theory would predict that finishing this paragraph would take *more* effort than it would in the absence of other competing activities. Attending a concert would provide entertainment, socialization, and fodder for future discussion. These predicted benefits could, therefore, outweigh the benefits of finishing this paragraph. Of course, one could argue, and rightfully so, that resisting the urge to attend a concert is comparable to keeping one's hand out of the cookie jar, *a la* Baumeister et al., (1998). However, the critical difference between Kurzban and colleagues' model and ego depletion is that the dynamic allocation of resources between the current task and other alternative tasks allows for situations where performance actually *improves* across time. If, instead of a concert, there were a polka dance festival occurring right now, which may appeal less to one's personal preferences, the perceived benefit of writing would remain largely unchanged, or perhaps even increase. Therefore, increases in performance resulting from, for example, a boost in positive affect (Tice et al., 2007), are not problematic for this model.

Promising though the theory proposed by Kurzban and colleagues may be, it is not without fault. A potential criticism could be that it is difficult to predict, *a priori*, the entire set of alternative activities that could arise during a given task. One could argue that often, alternative activities occur spontaneously and unpredictably; it would be difficult to predict whether or not a friend will decide to call, or that a boss will send an urgent email, during the course of a task, unless these events were planned ahead of time. How, then, can we make decisions such as which would be easier between two tasks, or which we would prefer to do first, without knowing the entire set of tasks in advance? An intuitive possibility is that, rather than relying on external features when predicting the relative effort of two tasks, such as potential alternative tasks, one might instead more directly compare features that the two tasks share in common. What exactly these shared features are is an open question addressed in the present thesis. More specifically, in Experiments 4-6 of the series of experiments to follow, I tested whether the shared variable used to compare the difficulty of cognitive and physical tasks is the difference in measured time (s) to complete either task.

CHAPTER 2. UNIMODAL COST COMPARISONS

In Experiments 1 and 2, I compared two physical costs, namely the cost of walking some distance versus reaching some distance. To do so, I married the design used to study pre-crastination and the design used by Rosenbaum (2008) in the study of walking and reaching. In the first two experiments, I asked participants to walk some distance and reach some distance to pick up and carry one of two buckets to a corresponding far table. In a typical critical trial, participants chose between, for example, reaching a longer distance to pick up and carry a bucket closer to their starting position, versus reaching a shorter distance to pick up and carry a bucket farther from their starting position to its corresponding far table. In Experiment 1, both buckets were empty. In Experiment 2, I added a lighter physical weight load (3.5 lbs.) to one of the buckets, and a heavier physical weight load (7.0 lbs.) to the other bucket.

Experiment 1: Added Reach; No Added Weight

The general walking and reaching paradigm is shown in Figure 1. Participants stood at one end of a 2-foot (0.61 m) wide alley bordered by waist-high string and faced two buckets. The buckets occupied waist-high tables to the left and to the right side of the alley. I varied two factors regarding the positions of the left and right buckets: the walking distance to either bucket, and the reaching distance to either bucket. Either of the two buckets could be located at 4 feet (1.22 m) or 12 feet (3.66 m) from the participant's start position. Additionally, either of the two buckets could be adjacent to the edge of the alley or 80 percent of the participant's arm length (M_{length} = 27.89 in.) away from the edge of the alley. Two additional tables stood to the left and the right at the end of the alley, 16 ft. (4.88 m) from the start position. I asked participants to walk down the alley in a natural way without stopping and to do whatever seemed easier: pick up the left bucket with their left hand and carry it to the far left table, or pick up the right bucket with their right hand and carry it to the far right table. I informed participants that there was no correct or incorrect answer, that I was simply interested in which of the options they found easier. However, the participants were instructed to try to avoid bumping into or knocking into the string boundary with his or her body while reaching for the bucket. If the participant knocked over the string boundary, s/he would have to repeat that trial.

The buckets were two bright yellow plastic beach pails, which were 5 in. (12.7 cm) high, with bases 4 in. (10 cm) in diameter and tops 7 in. (17.8 cm) in diameter. Each buckets stood on the center of its own circular wooden table, which were 24 in (0.12 m) high and 36 in. (0.91 m) in diameter. In each trial, each bucket's upright dark blue handle stood perpendicular to the long edge of the alley. The waist-high string boundaries lining the left and the right side of the alley were made of white cotton twine. Each string was held up by a single vertically placed .75 in. (1.91 cm) diameter plastic pipe at the start of the alley, and was attached to the inside edge of the far table at the end of the alley.

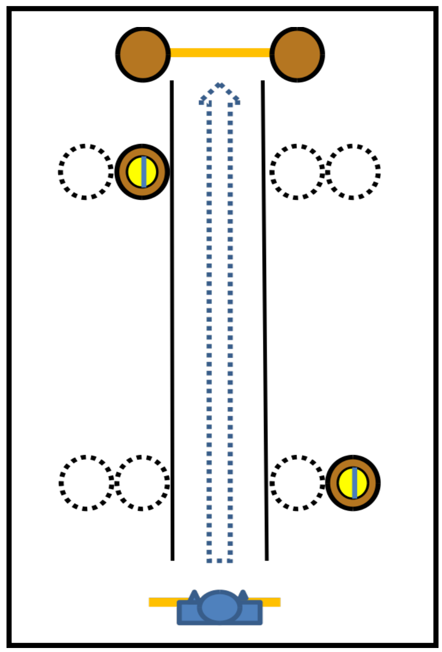

Figure 1. Diagram of the walking and reaching task.

Two experimenters were present for each experimental session. At the start of each session, participants sat in another room within the laboratory to fill out their informed consent as well as the Edinburgh handedness inventory (Oldfield, 1971). Once participants completed their informed consent, an experimenter asked for permission to measure his or her arm. After receiving verbal permission, the experimenter asked the participant to extend one of his or her arms at the start of the experimental session, and then measured approximately from the acromion (the bony protrusion extending over the shoulder joint) to the tip of the middle finger. The experimenter then calculated 80 percent of this value, which constituted the farthest possible reaching distance from the inside edge of the alley to the handle of the bucket. While the participant was still seated in the other room, both experimenters placed small strips of duct tape to the left and the right side of the alley to mark that participant's maximum reaching distance. After the strips of tape were correctly placed, the participant entered the main room of the laboratory where testing was to occur.

One end of the alley was marked with a bright green strip of horizontal duct tape. At the start of each trial, participants stood with the tips of their toes just behind this strip of duct tape, and roughly centered their body between the left and right string boundaries. Once the participants were in the correct location, they closed their eyes until an experimenter said the word "open." Prior to this, while participants stood with their eyes closed, both experimenters arranged the tables and buckets for the coming trial. When the buckets were in their correct location, one of the experimenters said the word "open." When the participant heard the word "open," s/he opened his or her eyes, and carried the bucket of his or her choice to its corresponding far table. Once s/he placed the bucket on its corresponding far table, s/he turned around, walked back to the start position, repositioned his or her feet on the starting line, and closed his or her eyes to await the next trial. After the participant completed all of the 16 trials, s/he was debriefed and dismissed.

Twenty-four participants were tested in all (22 Female, 2 Male; M_{age} = 19.42; Range = 19-20). Each participant completed a total of 16 trials (4 left bucket positions × 4 right bucket positions). 20 participants were right-handed, as determined by a score of 7 or higher on the Edinburgh handedness inventory.

Results: Experiment 1

In the original setting in which pre-crastination was tested, the buckets stood a short reach away from the edge of the alley. Therefore, it was possible that participants prioritized costs associated with walking because the costs associated with reaching were low. This possibility raised an interesting question, Would the tendency to show pre-crastination decrease if reaching costs were increased? To test this, I added long and short reaches to the paradigm in which pre-crastination was originally tested. Interestingly, despite the added reaching cost, participants continued to select the closer bucket in approximately 73% of trials tested. Said another way, in some trials, participants were willing to lean and reach far to pick up a bucket a shorter walk away.

Because participants continued to pick up the bucket a shorter walk away for the majority of trials, it was possible that they were simply not sensitive to the reaching manipulation. If this were the case, then a model including only the proportion of walking distance, as used in the pre-crastination series, should have accounted for the data just as well as a model including walking and reaching distance. Figure 2 shows the probability of selecting the right bucket as a function of an equation using only walking distance. The relevant equation was as follows:

$$p(\text{Right})=\text{AL}/(\text{AL}+\text{AR})$$

That is, the probability of choosing the right bucket, p(Right), was set equal to the walking distance (m) to the left bucket (AL), divided by the sum of that value and the walking distance (m) to the right bucket (AR). Said another way, it is the proportion of walking distance to the left bucket.

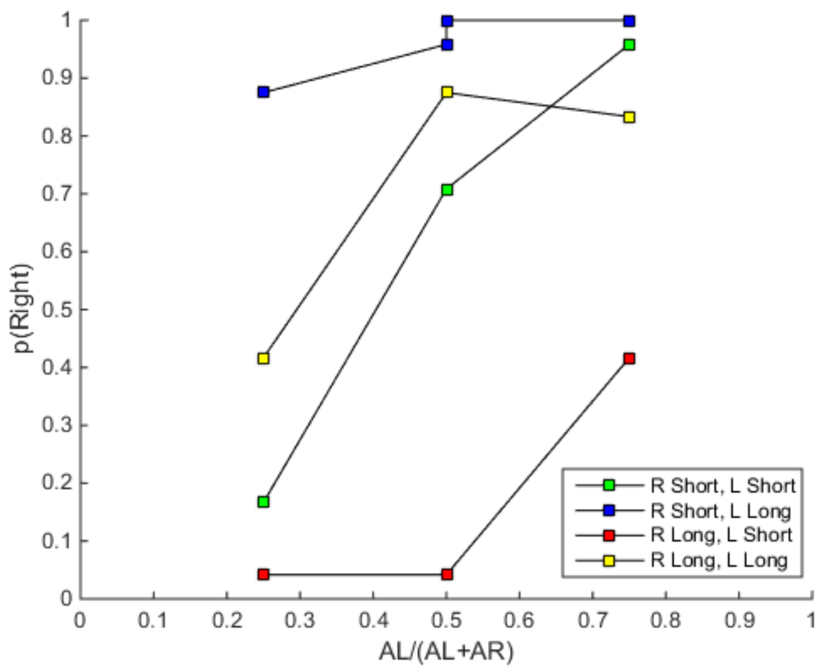

Figure 2. Probability of choosing the right bucket, or *p*(Right), plotted as a function of AL/(AL+AR). Blue squares show trials in which the right bucket stood adjacent to the edge of the alley, while the left bucket stood 80 percent of each participant arm lengths away from the edge of the alley. Yellow squares show trials in which both buckets stood 80 percent of participants' arm lengths away from the edge of the alley. Green squares show trials in which both buckets were adjacent to the edge of the alley. Red squares show cases in which the right bucket stood 80 percent of participants' arm lengths.

If participants were, in fact, sensitive to reaching cost, then the inclusion of reaching distance should reveal a more orderly pattern of data. Figure 3 shows the probability of choosing the right bucket, or *p*(Right) as a function of not only the proportion of walking distance, as shown in Figure 2, but also the proportion of reaching distance. The relevant equation used was as follows:

$$p(\text{Right}) = AL/(AL+AR) + RL/(RL+RR)$$

That is, the probability of selecting the right bucket was set equal to the sum of the proportion of walking distance to the left bucket and the proportion of reaching distance to the left bucket. RL is the reaching distance to the left bucket (m) from the inside of the alley, and RR is the reaching distance to the right bucket (m) from the inside of the alley. Fitting a logistic curve to *p*(Right) yielded a highly significant correlation, $r=.98$, $p<.001$.

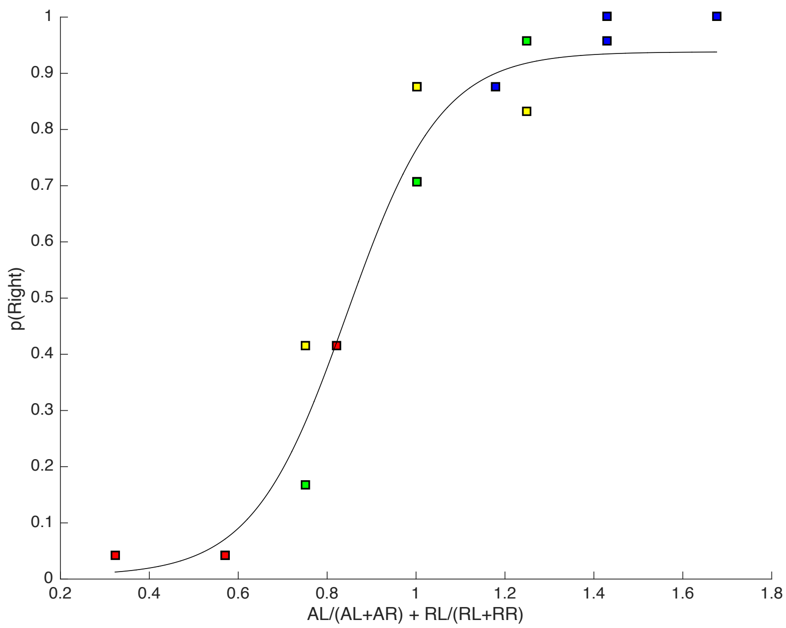

Figure 3. Probability of choosing to carry the right bucket, or *p*(Right), plotted as a function of AL/(AL+AR) + RL/(RL+RR), or the sum of the proportion of approach distance to the left bucket and the proportion of reaching distance to the left bucket.

Discussion: Experiment 1

In Experiment 1, I added long and short manual reaches to the design in which pre-crastination was originally tested. Therefore, a primary point of discussion relates to whether or not participants continued to show pre-crastination. Indeed, despite the added cost of reaching, participants continued to pick up the closer bucket, at least in terms of walking distance, for the majority of trials. However, the general pattern of data shown in Figure 3 does suggest some sensitivity to reaching, such that when there was a long reach to the right bucket and a short reach to the left bucket (red squares), the probability of choosing the right bucket was relatively low. Participants became slightly more willing to choose this bucket when it was only a short walk away, as shown by the rightmost red point. When there was a short reach to the right bucket and a far reach to the left bucket (blue squares), the probability of selecting the right bucket was nearly at ceiling. Therefore, although participants did continue to show pre-crastination, they were not oblivious to their reaching costs.

This finding raised the question of whether pre-crastination is rooted more in manipulation or locomotion. In the original pre-crastination experiments, I had no way to distinguish between these two factors. If pre-crastination is rooted more in locomotion, participants should have continued to pick up the bucket a shorter walk from their start position for the majority of trials. If pre-crastination is rooted more in manipulation, participants should have selected the bucket that was a shorter reach away from the inside of the alley. Participants continued to choose the bucket a shorter walk away, which supported the locomotion interpretation of pre-crastination.

A second point of discussion relates to the two models tested. The first model included only the proportion of walking distance to the left bucket. The clumping of data points and overall spread of the data suggested that some good variance (task-related) had not been accounted for. This is not surprising, as the model did not include reaching distance. When the proportion of reaching distance to the left bucket was also added, a very orderly S-shaped curve was revealed, yielding a strong logistic fit. Critically, the shape of the function is similar to that obtained in the original pre-crastination series, reinforcing the notion that the close-bucket preference persisted in spite of the added cost of reaching.

Experiment 2: Added Reach; Added Weight

In Experiment 1, I tested whether participants showed pre-crastination under costlier reaching conditions. To do so, I added long and short manual reaches to the paradigm in which pre-crastination was originally tested. In a typical critical trial, participants chose to either reach a longer distance to pick up and carry a bucket a shorter walk away, or reach a shorter distance to pick up and carry a bucket a longer walk away. Surprisingly, despite the added reaching cost, participants continued to pick up the bucket a shorter walk away for the majority of trials. This result suggested that participants continued to favor costs associated with *locomotion* rather than costs associated with *manipulation*.

Although participants continued to pick up the bucket a shorter walk away, it was not the case that they were oblivious to their reaching costs. As was previously mentioned, when there was a long reach to the right bucket and a short reach to the left bucket, the probability of picking up the right bucket was low. Conversely, when there was a short reach to the right bucket and a long reach to the left bucket, the probability of picking up the right bucket was high. This suggested that participants were sensitive to both their reaching *and* walking costs. Thus, it was still possible that under costlier manipulation conditions, participants would prefer to carry a bucket a shorter *reach* away rather than a bucket a shorter *walk* away.

In Experiment 2, I attempted to distinguish between the manipulation and locomotion accounts of pre-crastination by further increasing the costs associated with manipulation. To do so, I added 3.5 pounds (1.59 kg) of pennies to one of the two buckets, and 7 pounds (3.18 kg) of pennies the other bucket. For half the participants, the heavier bucket remained on the left side for all trials, and for the remaining half, the heavier bucket remained on the right side for all trials. Each of the two buckets was covered with a foam lid to occlude the pennies inside. For the first half of the participants, the heavier bucket was covered by a blue lid and the lighter bucket was covered by an orange lid. For the remaining half of the participants, the lids were switched to avoid any biases related to color. A secondary question of interest was to test whether the model that included only the proportion walking distance and the proportion of reaching distance to the left bucket used in Experiment 1 would predict participants' decisions concerning reaching for and carrying buckets that contained added weight loads.

The procedure used in Experiment 2 was very similar to Experiment 1. Two experimenters were present for each testing session. After one experimenter calculated the participants' arm length, they asked the participant to return to the room where testing would occur. In that room, a second experimenter had placed the two buckets on the two end tables at the end of the alley. An experimenter asked the participant to stand facing the two tables, pick up each of the two buckets, one at a time, and feel the weight they contained. Half the participants felt the weight of the left bucket first, and the remaining half felt the weight of the right bucket first. These groups were nested within each of the heavier-bucket-on-the-left and heavier-bucket-on-the-right groups of participants. Once the participant felt the weight of each bucket, s/he returned to the start of the alley, closed his or her eyes, and awaited the first trial. The procedure was identical to Experiment 1 for the remainder of the experiment.

Twenty-four participants were tested in all (21 Female, 3 Male; M_{age} = 19.04; Range = 18-21). Each participant completed a total of 16 trials (4 left bucket positions × 4 right bucket positions). All 24 participants were right-handed, as determined by a score of 7 or higher out of 10 on the Edinburgh handedness inventory.

Results: Experiment 2

In Experiment 2, I added light and heavy weights to the buckets to test whether the added cost associated with manipulation would cause participants to prefer the bucket a shorter *reach* away rather than a shorter *walk* away. Said another way, I hoped to distinguish between the manipulation and locomotion accounts of pre-crastination. Of secondary interest was testing whether a model including only the proportion of walking distance and the proportion of reaching distance to the left bucket, as used in Experiment 1, would continue to significantly predict the probability of selecting the right bucket.

To test whether participants continued to carry the bucket a shorter walk away from their start position, I measured the goodness of fit of the model used in the original pre-crastination series, which included only the proportion of walking distance to the left bucket. Figure 4 shows *p*(Right) as a function of AL/(AL+AR). Squares are conditions in which the heavier bucket was on the right side of the alley. Circles are conditions in which the heavier bucket was on the left side of the alley. The model yielded a poor correlation with *p*(Right), r=.02, p=.90. In fact, when tested in this way, the probability of carrying the bucket a shorter walk away fell to chance (approximately 51% of trials).

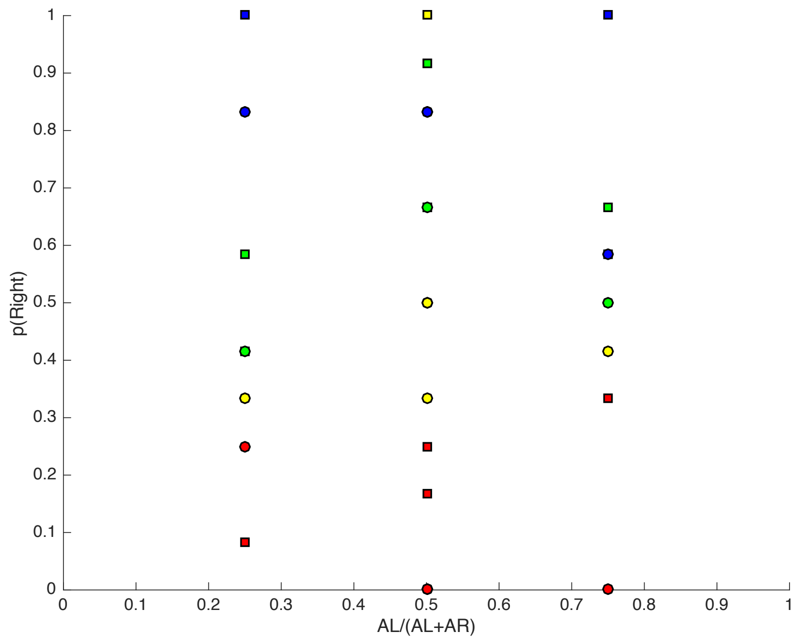

Figure 4. Probability of choosing the right bucket as a function of AL/(AL+AR). Squares represent conditions in which the heavier bucket remained on the left. Circles represent conditions in which the heavier bucket remained on the right. Blue squares show trials in which the right bucket stood adjacent to the edge of the alley, while the left bucket stood 80 percent of each participant arm lengths away from the edge of the alley. Yellow points show trials in which both buckets stood 80 percent of participants' arm lengths away from the edge of the alley. Green points show trials in which both buckets were adjacent to the edge of the alley. Red squares show cases in which the right bucket stood 80 percent of participants' arm lengths.

Next, I tested whether the model used in Experiment 1, which included both the proportion of walking distance to the left bucket and the proportion of reaching distance to the left bucket, would hold under conditions involving weighted objects. In other words, I used this model to test the null hypothesis, which was that the addition of light and heavy weights would not affect participants' decisions.

Figure 5 shows the probability of selecting the right bucket, or *p*(Right) plotted as a function of the sum of the proportion of walking distance and the proportion of reaching distance to the left bucket. Surprisingly, the model, which did not include the weight of either bucket, significantly predicted the obtained *p*(Right) values, $r=.75$, $p<.01$. This result may have coincided with the fact that there was no difference in the probability of selecting the right bucket for the groups with the heavier bucket on the left ($M_{percent}$ = 60.94) and the heavier bucket on the right ($M_{percent}$ = 43.67), $t\ (30) = 1.57$, $p = .26$.

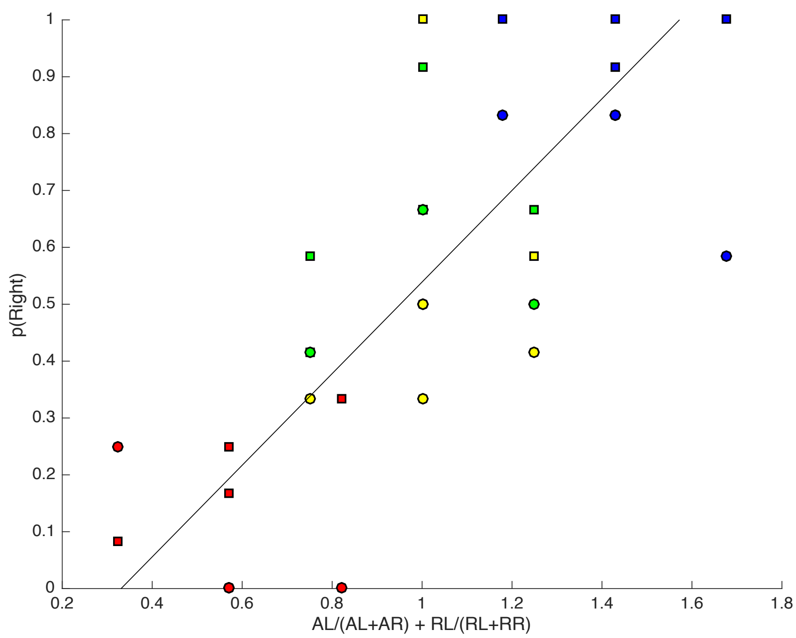

Figure 5. Probability of choosing the right bucket as a function of AL/(AL+AR) + RL/(RL+RR). Squares represent conditions in which the heavier bucket remained on the left. Circles represent conditions in which the heavier bucket remained on the right.

However, it is important to note that, although there was no difference in the probability of selecting the right bucket when it contained the lighter or the heavier weight load, it was unlikely that participants were oblivious to the added weight loads altogether. To test participants' sensitivity to the weight contained in either bucket, I used the following equation:

$$p(\text{Right})=\text{AL}/(\text{AL}+\text{AR}) + \text{RL}\times\text{WL}/(\text{RL}\times\text{WL}+\text{RR}\times\text{WR})$$

That is, the probability of selecting the bucket was set equal to the sum of the proportion of walking distance (m) to the left bucket and the proportion of the left reaching distance (RL) times the left bucket weight (WL) out of the sum of that value and the right reaching distance (RR) times the right bucket weight (WR). The model yielded a strong linear fit to the obtained *p*(Right) values across conditions, r=.80, p<.001, as shown in Figure 6.

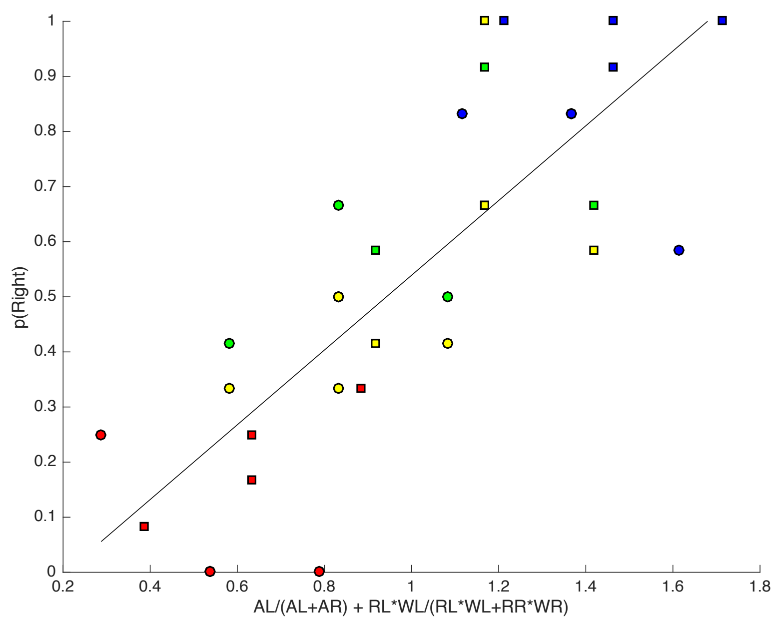

Figure 6. Probability of choosing the right bucket p(Right) as a function of AL/(AL+AR) + RL×WL/(RL×WL+RR×WR). Squares represent conditions in which the heavier bucket remained on the left. Circles represent conditions in which the heavier bucket remained on the right.

Finally, I wondered whether the lack of a close-bucket preference reflected a shift in participants' planning. Instead of planning only in terms of the walking distance to the bucket, participants may have also considered how far the bucket needed to be carried. To test this hypothesis, I added the proportion of left carry distance (CL) times the left bucket's weight (WL) out of the sum of that value and the right carry distance (CR) times right bucket's weight (WL), to the current model. Figure 7 shows *p*(Right) plotted as a function of the model just described. Indeed, the correlation with *p*(Right) increased numerically with the inclusion of carry distance, $r=.89$, $p<.001$.

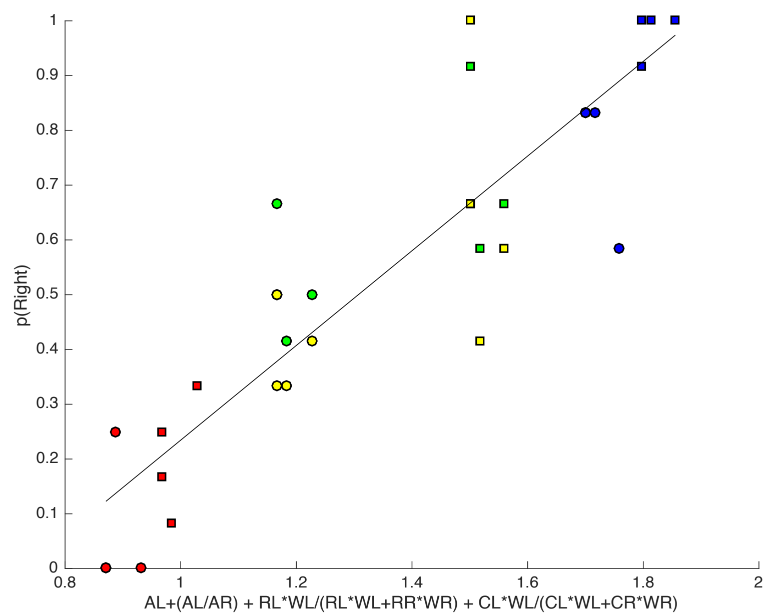

Figure 7. Probability of choosing the right bucket *p*(Right) as a function of AL/(AL+AR) + RL×WL/(RL×WL+RR×WR) + CL×WL/(CL×WL+CR×WR). Squares represent conditions in which the heavier bucket remained on the left. Circles represent conditions in which the heavier bucket remained on the right.

Discussion: Experiment 2

Despite the added reaching cost, participants tested in Experiment 1 continued to show pre-crastination in terms of locomotion. That is, they preferred to pick up the bucket a shorter *walk* away rather than a bucket a shorter *reach* away. However, participants were still sensitive to their reaching costs, suggesting that they considered costs associated with both locomotion *and* manipulation. Thus, the possibility remained that under costlier manipulation conditions, participants would no longer prefer to pick up the bucket a shorter walk away. To this possibility, I increased manipulation costs by adding heavy and light weights to the buckets. A secondary question concerning the weight manipulation was whether the model tested in Experiment 1, which included only the sum of the proportion of approach distance and the proportion of reach distance to the left bucket, would continue to significantly predict *p*(Right) when both buckets contained weight.

To address the first question concerning the locomotion and manipulation accounts of pre-crastination, when tested in this way, the probability of selecting the closer bucket dropped nearly to chance. This finding suggested that participants no longer favored the closer bucket, at least in terms of locomotion. Instead, they seemed to minimize costs associated manipulation, such as the distance to be reached for a bucket, the weight it contained, and how far it needed to be carried. This result is sensible in terms of minimizing physical risk. In Experiment 1, participants were occasionally willing to lean and reach far to pick up a bucket a short walk away. One could argue that this decision was physically risky, insofar as leaning and reaching far is a risk to one's balance. Adding weight to either bucket potentially increased this risk. Therefore, carefully considering manipulation costs, rather than relying solely on locomotion costs, could help to prevent potential injuries. This is especially true when the to-be-manipulated objects require far reaches or contain heavy weight loads.

However, it is also worth noting that the reduction in pre-crastination as previously discovered may have reflected an emerging desire to minimize time spent reaching for and carrying one of the two buckets. It is likely that participants needed to slow down their walking, at least to some extent, when reaching far to pick up a heavier bucket. This raises the possibility that, in some conditions, picking up the bucket a farther walk away actually took *less* time than picking up the bucket a shorter walk away. Consider conditions in which the heavier bucket was a shorter walk but a longer reach away, while the lighter bucket was a longer walk but a shorter reach away. Taking the few extra steps to pick up the farther, lighter bucket may have taken less time than slowing down to reach far for the closer, heavier bucket. Therefore, it is possible that, in conditions like these, walking a longer distance initially allowed participants to pick up the bucket more quickly. Such a finding would accord with pre-crastination, as it would suggest that participants continued to hasten the completion of a sub-goal, namely, grasping the handle of one of the two buckets.

To address the second question concerning the models tested, surprisingly, the model used in Experiment 1, which included only the sum of the proportions of walking and reaching distance significantly predicted the observed probabilities of selecting the right bucket in Experiment 2. This result was encouraging, as this very simple model including only physical values, rather than estimated coefficients, seems to be highly predictive of participants' selections even in situations involving objects with varied physical properties. However, two subsequent models, one of which added a term for weight and the other of which added a term for carry distance, suggested that participants were not *only* concerned with the initial walk and reach for the bucket. Instead, participants seemed to also be planning in terms of the distance that had to be walked with a heavy bucket in hand. To recall a previous point, it is possible that this shift in planning reflected a preference to reduce the degree to which participants had to slow down their walking speed, thereby minimizing time spent on the task. In Experiments 4-6 to follow, I will further discuss the relation between weight load and task time.

CHAPTER 3. MULTI-MODAL COST COMPARISONS

In the previous experiments, all of the choices were between two physical tasks. In Experiments 3-6, I compared cognitive and physical costs by asking participants to choose between performing a mainly cognitive task, namely counting aloud by ones to a specified number, and a mainly physical task, namely walking along and reaching some distance to pick up a bucket and carry it to a far table. In Experiment 3, I tested a metric to relate cognitive and physical costs. In Experiments 4-6, I tested the factor that participants used to make multi-modal decisions of this kind. More specifically, I tested whether participants used the difference in time between the two tasks to choose which to perform. In critical trials, participants chose, for example, to either do the bucket task or count aloud by ones at a leisurely pace to a low number. Said another way, in some trials, the count task took less time to perform than the bucket task, and vice versa. In Experiments 5 and 6, I further tested the time-based account of effort estimation by adding light

(Ex. 5; 3.5 lbs.) and heavy (Ex. 6; 7.0 lbs.) weight loads to the bucket. If participants considered *only* time when comparing between the two tasks, then the addition of weight should not have influenced participants' decisions: that is, unless it significantly increased the duration of the physical task. Due to the similarities in their procedures, the results of Experiments 4-6 will be discussed in tandem in the sections to follow.

Experiment 3: Count versus Carry: Long and Short Walks; Long and Short Reaches

The paradigm used in Experiment 3 is shown in Figure 8. The experimental setup was similar to that of the two previously described experiments, but with two major changes. The first change was that only one bucket was present in each trial. This bucket was either on the left or the right side of the alley. The walking and reaching distances to the bucket were identical to those used in the previous two experiments. The second change was that a computer monitor (32-in. Philips Model 32PFL4507/F7, Koninklijke Philips N.V., Amsterdam, The Netherlands) and a keyboard stood next to the participant's starting position. For half of the participants, this stand stood approximately 12 inches (30.48 cm) to the left of their starting position. For the remaining half of the participants, this stand stood approximately 12 inches (30.48 cm) to right of their start position. In each trial, a choice appeared on this computer monitor. Each choice consisted of one mainly physical task, namely carrying the bucket to its corresponding end table, and one mainly cognitive task, namely counting aloud by ones to either 10 or 50. Participants determined the rate at which to perform either task. For example, a typical choice was, "Would you rather carry the bucket to the far table (press the 'b' key) or count aloud to 10 (press the 'c' key)?" When the participant made his or her choice, s/he pressed either the 'b' key for 'bucket,' or the 'c' key for 'count,' followed by the 'Enter' key. Then, s/he performed the task s/he selected.

At the start of each trial, the participant was asked to attend to the computer screen, which displayed the words, "Please close your eyes until you hear the word 'open'. When you hear the word 'open,' please open your eyes and press the 'Enter' key." While the participant's eyes were closed, an experimenter moved the bucket to the correct location according to the specifications for that trial. When the experimenter had successfully placed the bucket in its correct location, s/he said 'open,' and the participant opened his or her eyes. Once the participant opened his or her eyes, s/he pressed the ‘Enter’ key to advance to the next screen, which displayed the pair of physical and cognitive tasks for that trial. After the participant made his or her selection and performed the task that s/he had selected, s/he pressed the ‘Enter’ key one more time, and the reminder to close his or her eyes appeared again.

24 participants were tested in all (17 Female, 7 Male; M_{age} = 19.29; Range = 18-22). Each participant completed a total of 16 trials (8 bucket positions × 2 levels of counting). All 24 participants were right-handed, as determined by a score of 7 or higher on the Edinburgh handedness inventory. All computer stimuli were generated in MATLAB.

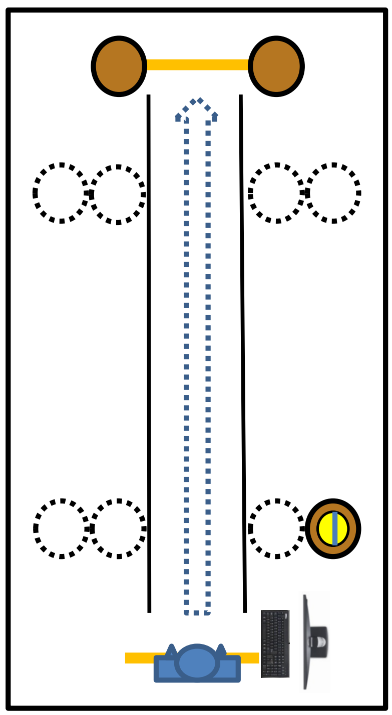

Figure 8. Schematic of Experiment 3. The participant (blue figure at bottom) used a keyboard to select whether s/he would like to carry the bucket (shown at only one of 8 possible locations) to its end table, or count aloud to either 10 or 50. Choices were displayed on the computer monitor to the participant’s right. Note that for half of the participants, the computer monitor and keyboard remained on the right, and for the other half of participants, the computer monitor and keyboard remained on the left.

Results: Experiment 3

In Experiment 3, I tested the cost of counting aloud, a mainly cognitive task, relative to walking per unit distance, a mainly physical task. To do this, I programmed an iterative algorithm in MATLAB to empirically estimate three coefficients: (1) the cost of reaching with the left hand relative to walking per unit distance (β_1); (2) the cost of reaching with the right hand relative to walking per unit distance (β_2); and (3) the cost of counting aloud each additional digit relative to walking per unit distance (β_3). The relevant equation was as follows:

$$p(\text{Bucket})=\text{CC}/(\text{PC}+\text{CC})$$

That is, the probability of carrying the bucket, or p(Bucket), was set to equal the proportion of cognitive cost out of the sum of the physical and cognitive costs, where physical costs (PC) were defined in the following way:

If the bucket stood on the left:

$$\text{PC}=\text{LWD} + (\text{LRD} \times \beta_1)$$

If the bucket stood on the right:

$$\text{PC}=\text{RWD} + (\text{RRD} \times \beta_2)$$

When the bucket stood on the left side of the alley, the physical cost was set equal to the sum of the walking distance (m) to the left bucket (LWD) and the product of the reaching distance (m) to the left bucket (LRD) and a coefficient, β_1. That coefficient represented the cost of reaching with the left hand relative to walking per unit distance. When the bucket stood on the right, the physical cost was set equal to the sum of the walking distance (m) to the right bucket (RWD) and the product of the reaching distance (m) to the right bucket (RRD) and a coefficient, β_2. That coefficient represented the cost of reaching with the right hand relative to walking per unit distance.

Cognitive cost was calculated by the number of digits to be counted (D) multiplied by the cost of counting aloud each additional digit relative to walking per unit distance (β_3). The relevant equation was as follows:

$$CC = D \times \beta_3$$

Figure 9 shows the probability of carrying the bucket plotted as a function of the proportion of cognitive cost out of the sum of the physical and cognitive costs, as well as the associated coefficients. As a whole, the predicted probabilities generated using the model above and the observed probabilities of carrying the buckets were highly correlated, r=.93, p<.001. This correlation was achieved when the right hand (β_1) and left hand (β_2) reaching costs were set to 4, and the cost of counting aloud each additional unit (β_3) was set to 32. Said another way, reaching per unit distance with the right and left hand was four times costlier than walking per unit distance, and counting aloud each additional digit was 32 times costlier than walking per unit distance. Moreover, the cost of counting aloud each additional digit was approximately eight times costlier than reaching an additional unit of distance.

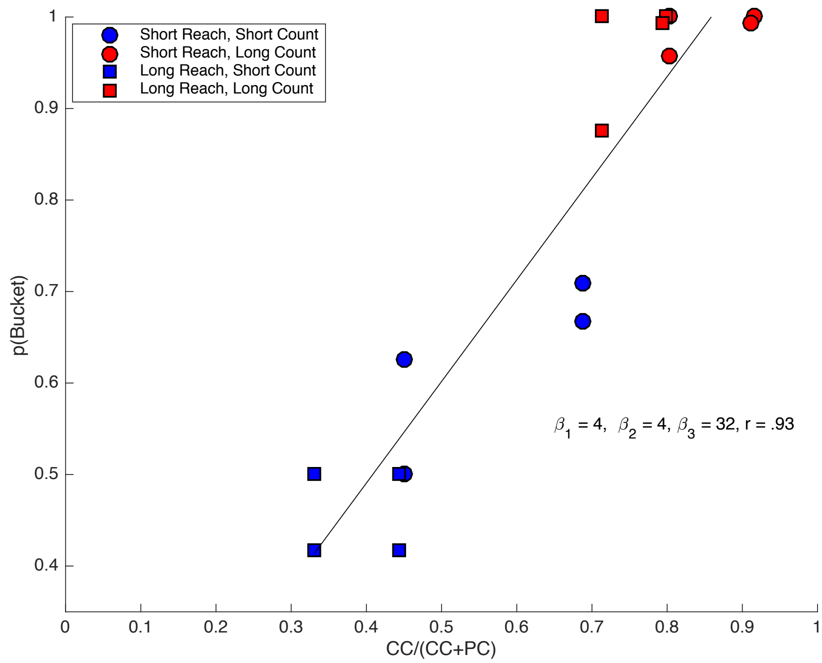

Figure 9. Probability of carrying the bucket plotted as a function of the proportion of cognitive cost out of the sum of cognitive and physical costs. Blue circles are conditions where participants chose between counting aloud to 10 and reaching a short distance for a bucket. Red circles are conditions where participants chose between counting aloud to 50 and reaching a short distance for a bucket. Blue squares are conditions where participants chose between counting aloud to 10 and reaching 80 percent of their arm length for a bucket. Red squares are conditions where participants chose between counting aloud to 50 and reaching 80% of their arm length for a bucket. Note that for aesthetic purposes, the 'y' axis is constrained between 3.5 and 1.

Discussion: Experiment 3

In Experiment 3, I tested the relative costs of performing a mainly cognitive task, namely counting aloud, and a mainly physical task, namely walking some distance and reaching some distance to pick up a bucket and carry it to a corresponding end table. There are four main points of discussion concerning the method and results of Experiment 3.

A first point of discussion concerns the estimated coefficients obtained in Experiment 3. The strongest correlation between the observed and estimated probabilities of choosing to carry the bucket was achieved when both right and left reaching costs were set to 4 times that of walking per unit distance, and when the cost of counting each additional unit was set to 32 times that of walking per unit distance. These coefficients suggest that cognitive cost "won the day," so to speak. Counting aloud seemed to be so costly, in fact, that the costs of reaching with the left and right hand were less than half as costly as previously calculated costs of reaching (Rosenbaum, 2008). This result is sensible when one considers the rather large 'bucket bias,' as shown in Figure 9. The probabilities of selecting the bucket started at approximately 40 percent and rose steeply. If participants were trying hard to avoid counting aloud, they may have become more willing to incur physical costs to do so, accounting for the reduction in reaching coefficients. The asymmetry between cognitive and physical costs supports the working memory account of pre-crastination, in which pre-crastination is thought to reflect a desire to reduce cognitive cost at the expense of physical effort.

A second point of discussion concerns the validity of the coefficient representing cognitive cost. Note that I have refrained from discussing which aspect of counting was most related to cost; I have simply referred to an overall cost of the task itself. I did so intentionally in recognition of the fact that this cost may have been a byproduct of any number of factors. These factors may have been directly related to more cognitive aspects of counting, such as tracking one's own progress throughout the series of digits. However, it is also possible that these factors were more social in nature. Some participants claimed that they felt more awkward standing idle and counting aloud than they did carrying a bucket, though both tasks were being observed in the same way. Thus, though the coefficients suggested that I captured some kind of cost associated with counting, what exactly that cost was remains unclear. Moreover, although I chose to fit a linear model to the data, it is possible that, like in Experiment 1, a logistic function or other non-linear function would have provided a significantly better fit. In doing so, it is admittedly possible that the values of the coefficients estimated here would change.

A third point of discussion concerns how exactly participants were able to choose between these very different tasks. At the end of each experimental session, I asked participants what they were using to compare between the two tasks, and the majority responded by saying that they estimated the time that either task would take, and then picked the shorter task. This response was especially interesting, as it was a clever way to solve the "apples-and-oranges" problem they were presented with. Though purely speculation, these self-reports opened the door to the possibility that time, or more specifically the difference in measured time between the two tasks, is the substrate linking cognitive and physical effort. Such a notion accords with previous work showing that strategies are selected to minimize time, even when the difference in duration between strategies is on the order of milliseconds (Gray & Boehm-Davis, 2000). Even more exciting is the possibility that perceived time, as opposed to measured time, functions as a combined variable to allow us to judge costs across very different kinds of tasks. It is possible that the predicted effort of either task influenced the predicted time to complete the task, such that higher levels of predicted effort would result in longer judgments of time and lower levels of predicted effort would result in shorter judgments of time. In this way, predicted times could encompass effort. These ideas will be discussed at greater length in the General Discussion section to follow. However, it is equally important to note that, although I did not specify the rate at which participants should complete either task, participants tended to count at rather slow paces. In fact, some participants reported thinking about each digit counted as approximately one second. This feature may have biased participants to use time as a common substrate for the two tasks.

The fourth and final point of discussion involves issues with the stimuli used in the current design. I reasoned that, because it was unlikely that participants would not be able to reach for and carry the bucket, even at its most distant location, it was important to choose a cognitive task that was relatively easy to perform. Counting seemed to meet this criterion while still requiring a degree of cognitive skill. Because I did not specify the rate of the physical task, as previously mentioned, I also did not specify the rate of the cognitive task. Therefore, a priori, though I sought two levels of counting difficulty, these levels were difficult to choose without knowledge of the rate at which participants would count, the time it would take them to complete the bucket task, or how sensitive participants would be to their physical effort relative to their cognitive effort. In hindsight, knowing that participants counted at relatively slow paces, and knowing that little weight was given to the physical aspects of the task, counting aloud to 50 seems unnecessarily lengthy and difficult relative to the bucket task. Though this issue challenges the validity of the coefficients estimated here, it will inform the selection of future stimuli in tests of multi-modal effort estimation.

Experiments 4-6: Count versus Carry: Long and Short Reaches; Varied Weight Loads

Drawing upon participants' self-reports in Experiment 3, in Experiments 4-6, I tested more directly whether measured time is, in fact, the variable used to make multi-modal decisions of this kind. To do so, I capitalized on task times obtained in Experiment 3. Following the key press of either 'c' for count or 'b' for bucket, the participant pressed 'Enter' immediately before beginning the task s/he chose. This key press started a timer function in MATLAB. When the participant finished his or her task, s/he pressed 'Enter' again, which stopped the timer and recorded his or her task time. Using these task times, I calculated participants' mean count rate (.85 digits/s) across both count conditions (10 and 50 digits), as well as participants' average bucket task time (11.9 s). Then, I calculated the number of digits that participants would have had to count, at the rate obtained from Experiment 3, for the count task time to equal the bucket task time. This value turned out to be 14 digits. Using 14 digits as a median value, I chose two count values that I expected to take less time and two count values that I expected to take more time than the average bucket task time: 8, 12, 16 and 20 digits. The prediction was clear: if participants use time to compare between cognitive and physical tasks, then they should choose to count to 8 and 12 more frequently than they choose to carry the bucket, and likewise, should choose to carry the bucket more frequently than they choose to count to 16 and 20.

For an additional test of the time-based account of effort estimation, I reasoned that, if participants relied solely on the difference in time between the two tasks to decide which task to perform, then manipulating the demands of either task should not significantly influence their choices if the task time was not strongly affected. Therefore, I sought a manipulation that would change the difficulty of the physical task but not strongly influence the time. To test this hypothesis, in Experiments 5 and 6, I added weights to the bucket. Specifically, in Experiment 5, I added 3.5 lbs. of pennies, and in Experiment 6, I added 7 lbs. of pennies. Before testing began, I asked each participant to pick up the bucket with their right hand and their left hand to feel its weight. Half of participants picked up the bucket with their left hand first and then their right hand, and the remaining half picked up the bucket with their right hand first and then their left hand. To prevent any biases related to the visual appearance of the pennies the bucket contained, a dark blue foam lid concealed the contents of the bucket.

The procedure used in Experiments 4-6 was nearly identical to that of Experiment 3. In each trial, participants chose to either count aloud by ones at a leisurely pace to a specified value, or walk along, reach some distance to pick up a bucket, and carry it to a corresponding far table. All instructions and stimuli were presented in the same way as in Experiment 3. However, unlike in Experiment 3, where both the walking and reaching distances to the bucket varied, in Experiment 4, only the reaching distance to the bucket varied. The bucket always stood 8 ft. (2.44 m) from the participants' start position, as shown in Figure 10. As in all previous experiments, the maximum reaching distance was 80 percent of a given participant's maximum reach from the inside of the alley. Thus, there were four levels to the physical task: (1) right short reaches; (2) left short reaches; (3) right long reaches; and (4) left long reaches. All possibilities of the four count levels and four reach levels were tested, resulting in a total of 16 trials for each participant. 24 participants were tested in each experiment for a total of 72 participants (Ex. 4: 19 Female, 5 Male; M_{age} = 19.83; Range = 18-34; Ex. 5: 17 Female, 7 Male; M_{age} = 19.33; Range = 18-22; Ex. 6: 18 Female, 6 Male; M_{age} = 19.13; Range = 18-21;)[1]. 66 of the 72 participants tested were right-handed, as determined by a score of 7 or higher on the Edinburgh handedness inventory. Two participants in each of the three experiments were left-handed, as determined by a score of 7 or lower.

[1] Half of the participants in Experiment 6 were tested on the day of THON (2/16/16), which is a large student-organized dance marathon that occurs each year at Pennsylvania State University to raise money for pediatric cancer. Anecdotally, participants tested on the day of THON were unusually distracted and erratic in their decisions, as compared to the participants tested in Experiments 4 and 5. In an effort to rule out the possibility that the results of Experiment 6 were due to a sampling issue, I re-tested 12 participants to replace those participants tested on the day of THON. A comparison between two mixed model ANOVAs, one performed with the 12 participants tested on the day of THON and one performed with 12 new participants, revealed no differences in the statistical outcome between the two groups. To ensure that the results of Experiment 6 were comparable to Experiments 4 and 5, when students were not, presumably, drawn to participate in a large-scale University-wide event, I will default to the sample that does not include participants tested on the day of THON.

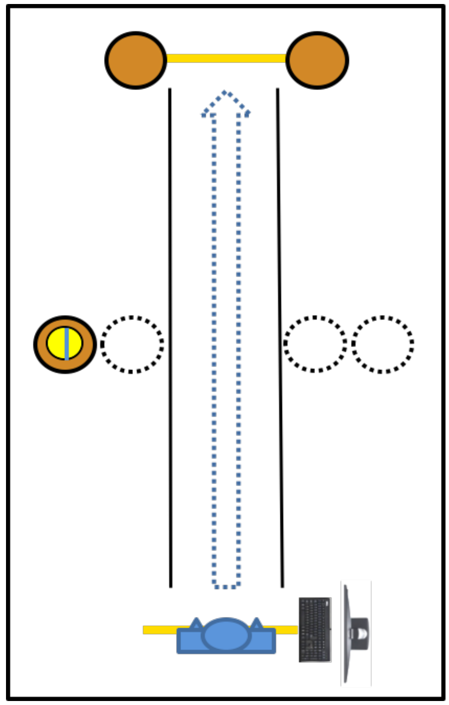

Figure 10. Schematic of the setup used in Experiments 4-6. The participant (blue figure at bottom) used a keyboard to select whether s/he would like to carry the bucket (shown at only one of 4 possible locations) to its end table, or count aloud to 8, 12, 16, or 20. Choices were displayed on the computer monitor to the participant's right. Note that when the bucket contained weight, as it did in Experiment 5 (3.5 lbs. of pennies) and Experiment 6 (7.0 lbs. of pennies) a blue foam lid (not shown here) occluded participants' view of the pennies it contained.

Results: Experiment 4-6

The results of Experiments 4-6 are shown in Figures 11-13, where *p*(Bucket) is plotted as a function of the four count values tested for Experiments 4-6, respectively. Red circles are conditions in which there was a short right reach to the bucket. Blue circles are conditions in which there was a short left reach to the bucket. Red squares are conditions in which there was a long right reach to the bucket. Blue squares are conditions in which there was a long left reach to the bucket. Across the three plots, one can see a strong and roughly linear increase of *p*(Bucket) across the four levels of count. Notably, participants seemed to discriminate much more heavily between short and long reaches across all levels of count when comparing between the no added weight (Experiment 4; Figure 11) and 3.5 lbs. of added weight (Experiment 5; Figure 12) conditions. Although the overall *p*(Bucket) values decreased, strangely, participants seemed *less* sensitive to reaching costs in Experiment 6 (Figure 13), where 7.0 lbs. of weight were added to the bucket. In fact, of the three experiments, participants seemed to discriminate the least between long and short reaches when paired with low count values in Experiment 6. These surprising results with be discussed at length in the Discussion section to follow.

A mixed model ANOVA was performed on Experiments 4-6 together with three within-subject factors (Hand: left, right × Reach: short, long × Count: 8, 12, 16, 20 digits) and one between-subject factor (Added Weight: 0 lbs., 3.5 lbs., 7 lbs.) was performed on the mean probability of selecting the bucket, *p*(Bucket). The ANOVA revealed a highly significant main effect of reach, $F(1, 69)=53.70$, $p<.001$, such that participants were significantly less likely to choose to carry the bucket when it required a long reach ($m=.45$) than when it required a short reach ($m=.70$). Additionally, there was a highly significant main effect of count, $F(2.33, 160.96)=56.09$, $p<.001$, such that *p*(Bucket) increased across the 8 digit ($m=.30$), 12 digit ($m=.51$), 16 digit ($m=.68$) and 20 digit ($m=.81$) conditions. The ANOVA revealed no other significant main effects or interactions. However, the weight × count interaction approached significance, $F(6, 136)=1.89$, $p=.08$, such that participants preferred to count more often as the weight of the bucket increased, or said another way, preferred to carry the bucket less as the weight of the bucket increased.

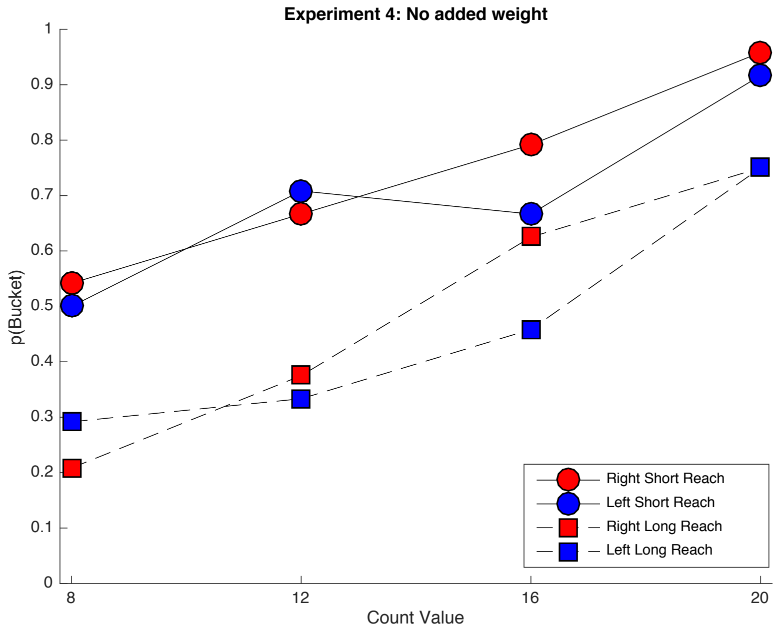

Figure 11. Probability of carrying the bucket, *p*(Bucket), plotted as a function of count value for Experiment 4. Red circles are conditions in which the bucket task required a short right reach. Blue circles are conditions in which the bucket task required a short left reach. Red squares are conditions in which the bucket task required a long right reach. Blue squares are conditions in which the bucket task required a long left reach.

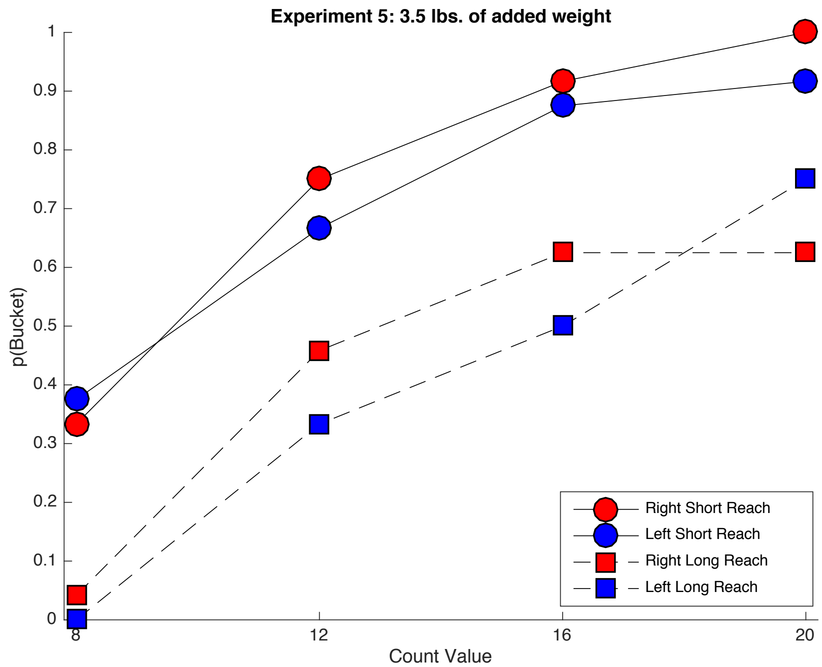

Figure 12. Probability of carrying the bucket, *p*(Bucket), plotted as a function of count value for Experiment 5. Red circles are conditions in which the bucket task required a short right reach. Blue circles are conditions in which the bucket task required a short left reach. Red squares are conditions in which the bucket task required a long right reach. Blue squares are conditions in which the bucket task required a long left reach.

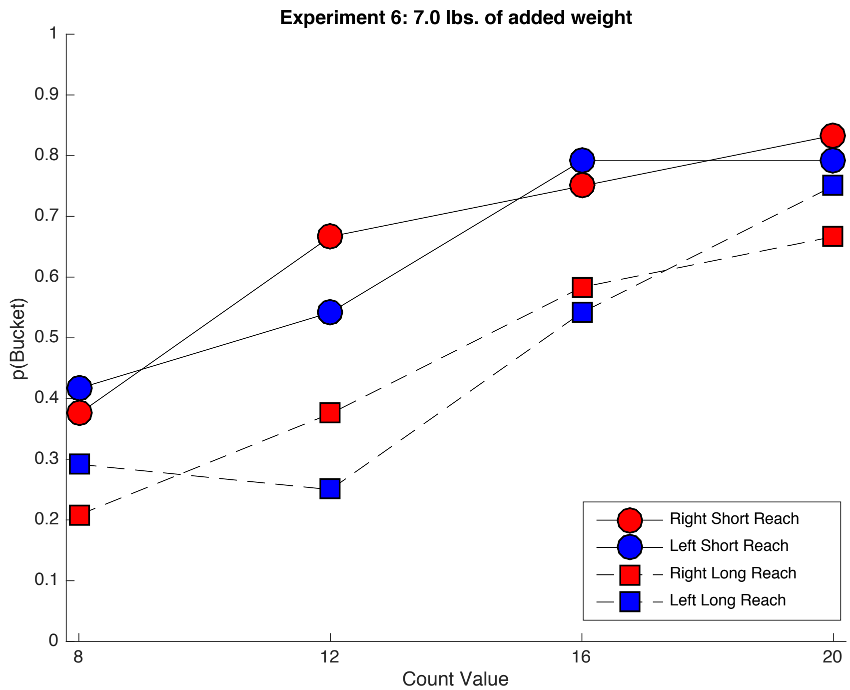

Figure 13. Probability of carrying the bucket, *p*(Bucket), plotted as a function of count value for Experiment 6. Red circles are conditions in which the bucket task required a short right reach. Blue circles are conditions in which the bucket task required a short left reach. Red squares are conditions in which the bucket task required a long right reach. Blue squares are conditions in which the bucket task required a long left reach.

In addition to the general analysis just described, a central topic of interest was the extent to which participants used measures time (s) as a way to compare between the two kinds of tasks. In pursuit of this topic, I was led to two questions: (1) Did the difference in time between the count and bucket tasks significantly predict *p*(Bucket), and (2) if so, did the correlation between the difference in time and *p*(Bucket) change as a function of physical weight load? To address these questions, I used task times obtained from Experiments 4-6. These task times were obtained via a timer programmed in MATLAB identical to the program used in Experiment 3. Indeed, the difference between task times significantly correlated with *p*(Bucket) across the 16 conditions tested, r=.76, p<.001, suggesting that participants were, in fact, using the difference in task time to inform their choices. Moving on to the second of the two questions just mentioned, Does the weight of the bucket impact this correlation: Interestingly, the difference in obtained task times continued to significantly predict *p*(Bucket) in both Experiment 5, r=.73, p<.01, and Experiment 6, r=.72, p<.01, in which the bucket contained 3.5 lbs. and 7 lbs. of weight, respectively. However, there was a slight though statistically negligible numerical decline in the associated r value across the three levels of weight.

As a complimentary analysis, I used the obtained count rates and mean bucket times from Experiments 4-6 to predict the point of subjective equality (PSE) that one would expect if time were the sole factor of interest for each experiment. Said another way, I calculated the number of digits that participants would have had to count, at the obtained count rate, for the count task time to equal the bucket task time. Specifically, I was interested in whether: (1) the predicted PSE would closely match the obtained PSE; and (2) the difference between the predicted and obtained PSE would grow as a function of physical weight load. Figures 14-16 show *p*(Bucket) collapsed across the four reach conditions plotted as a function of count value for Experiments 4-6, respectively. The predicted PSE's are shown in red, and the obtained PSE's are shown in green. To answer both questions simultaneously, the difference between the predicted PSE and obtained PSE was relatively small when the bucket contained no weight Experiment 4 (1.6 digits), and then grew slightly across Experiment 5 (2.7 digits) and Experiment 6 (3.9 digits). These results are consistent with the notion that, although the effect of weight was not statistically significant, participants may have relied on judgments of relative time slightly less as task demands increased.

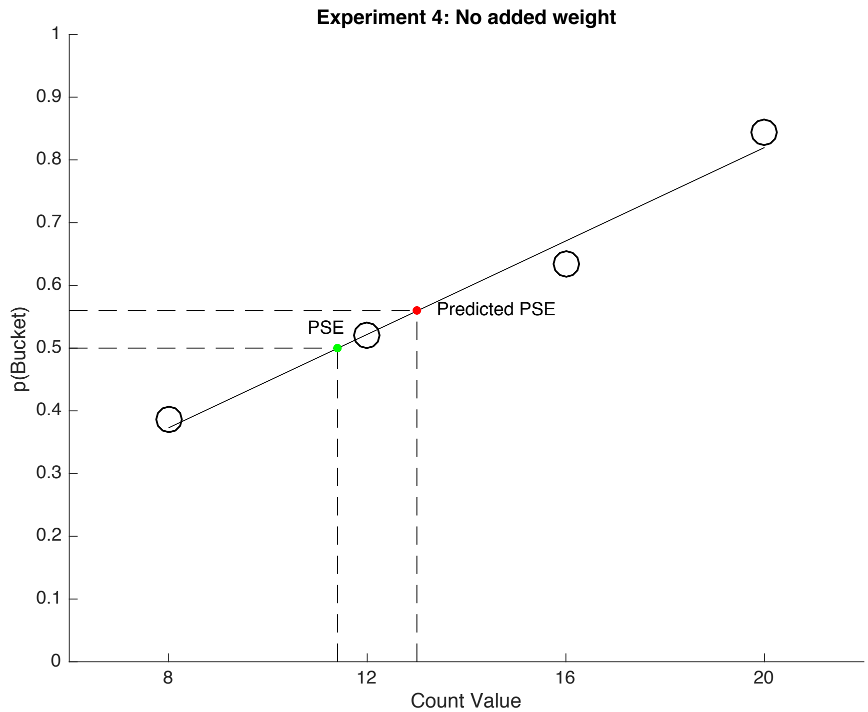

Figure 14. Probability of carrying the bucket collapsed across reach conditions plotted as a function of the number of digits to be counted in Experiment 4. The predicted point of subjective equality (PSE) is shown in red, while the obtained PSE is shown in green.

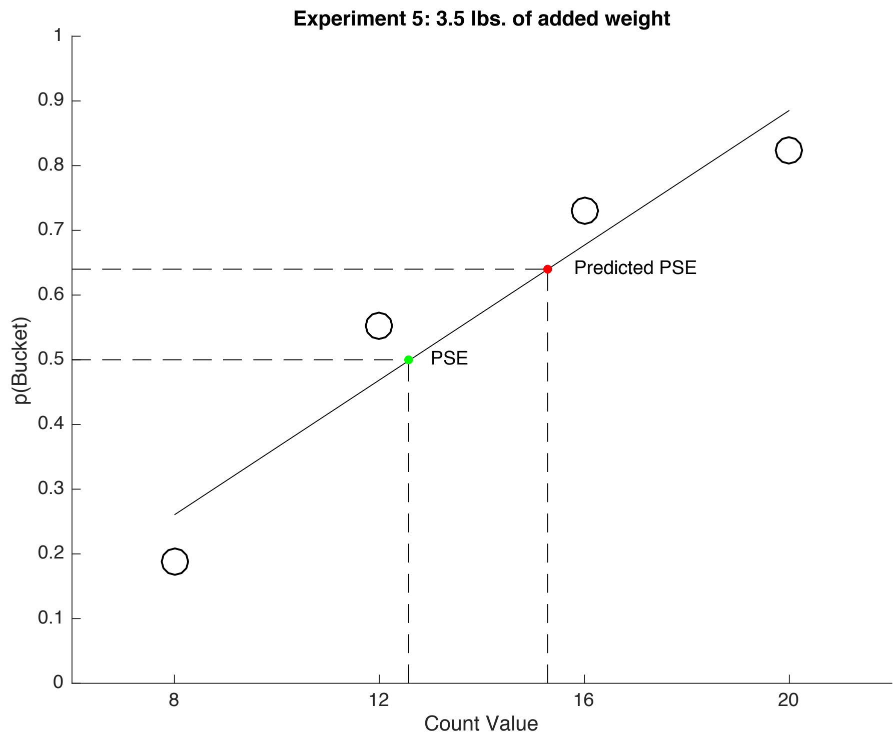

Figure 15. Probability of carrying the bucket collapsed across reach conditions plotted as a function of the number of digits to be counted in Experiment 5. The predicted point of subjective equality (PSE) is shown in red, while the obtained PSE is shown in green.

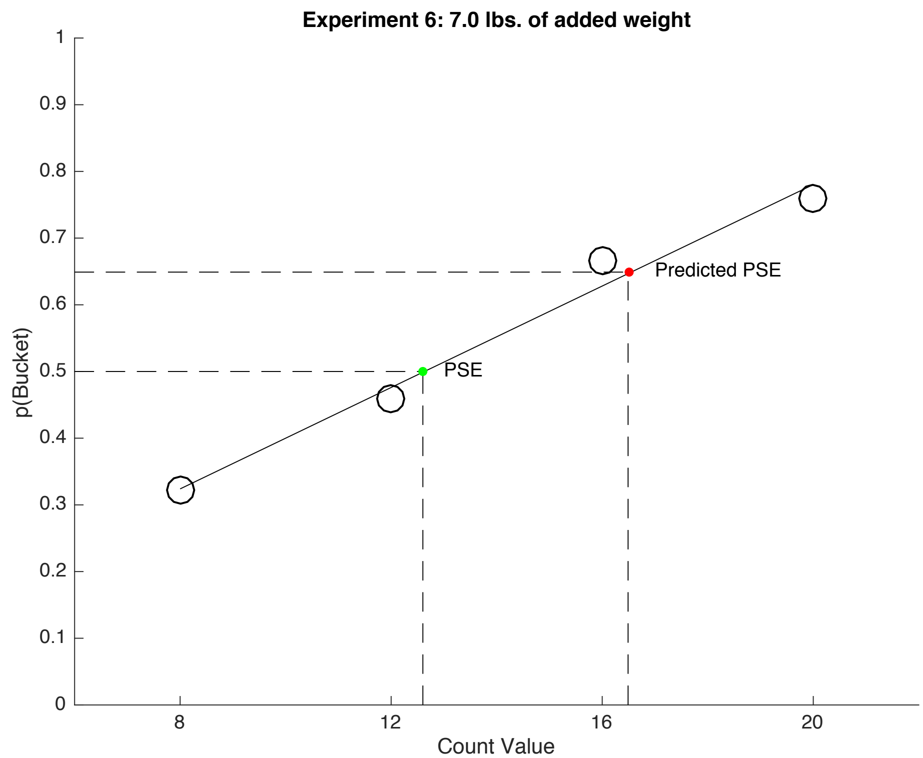

Figure 16. Probability of carrying the bucket collapsed across reach conditions plotted as a function of the number of digits to be counted in Experiment 6. The predicted point of subjective equality (PSE) is shown in red, while the obtained PSE is shown in green.

An additional question of interest is how the various count rates and task times compared across the three experiments. Interestingly, although the mean bucket task times increased slightly across Experiment 4 (11.23 s), Experiment 5 (12.17 s), and Experiment 6 (12.80 s) while mean count rates decreased across the three experiments (.86 digits/s, .80 digits/s, & .77 digits/s, respectively). Said another way, participants tended to count, on average, at slower rates when the bucket task took a longer time.

A final question of interest was whether, when tested on an individual level, the differences in task time continued to predict *p*(Bucket). However, an analysis concerning differences in obtained task times was impossible at an individual level; when a given participant chose to either count or carry the bucket, by nature of the task design, a task time was recorded for only one of the two tasks. If, for example, a given participant always chose to carry the bucket when it was paired with 20, then the recorded task time for counting to 20 would be missing from their data. Another issue was that task times depended on which tasks participants chose to perform. If a given participant only counted to low numbers, such as 8 or 12, then their mean count task time would be lower. If instead, they chose frequently to count to higher numbers, such as 16 or 20, then their mean count task time would be higher.

To circumvent these potential issues, I estimated task rates for participants who showed variation in their choices. If a given participant always counted or always counted the bucket, then they were omitted from the analysis. There were two such participants in Experiment 4, three in Experiment 5, and four in Experiment 6. Interestingly, with the exception of one participant in Experiment 4, all of these participants chose to carry the bucket in every trial. Task rates were calculated by dividing the task "distance," as it were, by the task time. For the bucket task, the relevant distance was the walking distance in centimeters. For the count task, the relevant distance was the number of digits counted. Surprisingly, neither participants' count task rates nor bucket task rates significantly predicted *p*(Bucket) in Experiments 4-6. (all p values > .05). However, this result is sensible when one considers that participants' bucket and count task rates also did not correlate (all p values >.05). Said another way, participants who counted quickly did not necessarily perform the bucket task quickly or slowly; there was a great deal of variation in the task rates.

It was possible that participants' task rates did not predict their choices because what mattered was not necessarily the rate at which they were able to complete either task individually, but rather the difference in total time between the two tasks, as was used in all previous analyses. Pursuing this question reinforced a problem that was already mentioned: many task times were missing from a given participants' data. In an attempt to obviate the prior collection of task times, I drew upon each participant's mean count task rate and mean bucket task rate to predict mean count and bucket task times. To do this, I calculated the product of each participant's mean count rate and the mean value of the four count conditions (14 digits), as well as the product of each participant's mean bucket task rate and the distance walked in each bucket task condition

(487.68 cm). Then, I tested whether the difference between these predicted task times significantly predicted the frequency with which a given participant chose to carry the bucket. In Experiments 4 and 6, the correlation between the difference in predicted task times and the frequency of selecting the bucket approached significance (Ex. 4: $r=.44$, $p=.05$; Ex. 6: $r=.78$, $p=.06$). However, in Experiment 5, where the bucket was filled with 3.5 lbs. of weight, the differences in the predicted task times yielded a rather poor correlation with the frequency of selecting the bucket across participants, $r=.12$, $p<.05$. This finding will be discussed in the section to follow.

Discussion: Experiments 4-6

In Experiments 4-6, I tested more directly whether participants used the difference between the total bucket and count task times to decide whether to count or carry the bucket. To do so, I drew upon task times obtained in Experiment 3 to estimate two count values that I expected to take less time than the bucket task and two count values that I expected to take more time than the bucket task. Additionally, I reasoned that, if participants relied solely on time to choose between cognitive and physical tasks, then increasing the difficulty of the physical task by introducing physical weight loads should not have influenced participants' decisions, assuming that the task time did not significantly change. To test this hypothesis, I added 3.5 lbs. of weight in Experiment 5 and 7 lbs. of weight in Experiment 6. Note that in Experiment 4, the bucket did not contain any added weight.

Though participants preferred to pick up the bucket less when it required a long reach, and likewise, preferred to count less as the number of digits to be counted increased, surprisingly these were the only effects to reach statistical significance. In other words, participants were just as willing to reach a long distance for a bucket that was filled with 7 lbs. of weight as they were for a bucket that contained no added weight. Likewise, though the majority of participants were right-handed, they were just as willing to heft a heavy bucket with their left hand as they were with their right. Most surprising was the fact that, across the three experiments, each with increasing physical weight values, the probability of choosing to carry the bucket did not significantly change. In fact, participants seemed to discriminate *less* between long and short reaches at low count values when the bucket contained the *most* weight (7.0 lbs.), as shown in Figure 13. These surprising findings call to mind the idea put forth by Fiske and Taylor (1991) that we tend to be "cognitive misers." That is, we tend to sacrifice optimality or accuracy to save precious cognitive resources. Though this term is generally used to refer to the quick-and-dirty heuristics that may give rise to phenomena such as stereotyping (Macrae et al., 1994), the general idea accords with the present findings as well as our interpretation of pre-crastination. If one allows that participants more heavily prioritized their cognitive effort than their physical effort, as they appeared to in Experiment 3, then the aforementioned results are sensible. However, the fact that the mean *p*(Bucket) values decreased numerically across levels of weight suggested that, if the bucket contained a sufficiently heavy weight load, participants would pick it up less. It is difficult to imagine, though admittedly possible, that participants would be willing to heft, for example, 100 lbs. of weight to avoid counting to 30 at a leisurely pace.

Regarding the time hypothesis, participants did, indeed, seem to use the difference in measured time between the two tasks to inform their decisions. The differences between the bucket task times and count task times yielded significant correlations with *p*(Bucket) across all three experiments. Numerically, these correlations decreased slightly across Experiments 4-6, though the decrease was statistically negligible. These results suggested that the participants' self report in Experiment 3 was correct; across Experiments 4-6 participants tended to use the differences in time between the two tasks to choose which task would be easier. However, the discrepancies between the predicted PSE's, which were calculated using only time, and the obtained PSE's grew across the three experiments. This growing discrepancy suggested that participants may have relied slightly less on differences in task time as task demands, namely the weight of the bucket, increased. Abandoning a time-based strategy to compare between cognitive and physical costs is sensible when one considers more extreme scenarios, and scenarios that are well beyond the parameters that would be considered safe for our participants. Consider the previously mentioned example in which the choice is between lifting 100 lbs. of weight and counting aloud to 30. Though lifting 100 lbs. of weight would likely take less time than counting aloud to 30, it is difficult to imagine that it would *feel* subjectively easier. At some point, task demands must also be incorporated into one's decisions to avoid sub-optimal behavior.

A final point of discussion is how to interpret the relatively strange pattern of results regarding individual differences in the use of time. To reinforce these results, across Experiments 4-6, neither a given participants' count task rate nor bucket task rate predicted the frequency with which they chose to carry the bucket. This result is sensible when one considers that the count task rates and bucket task rates were not predictive of one another. Counting quickly did not predict fast or slow bucket task rates, and walking quickly did not predict fast or slow counting rates. Therefore, it would have been surprising if participants would have relied *only* on the rate at which one task or the other could have been performed when choosing between the two tasks. Instead, a factor like the difference between the two total task times would take into account the rate of completion for *both* tasks, rather than just one. However, as previously noted, not all task times were recorded for all participants. Therefore, drawing upon each participant's bucket and count task rates, I estimated predicted task times, and correlated the differences between those predicted task times with the frequency of selecting the bucket task. Though these correlations were just shy of statistical significance in Experiments 5 and 6, the correlation was much weaker in Experiment 4. One could, therefore, surmise that the use of measured time as a variable to relate cognitive and physical effort followed a quadratic shape across levels of weight. However, it is important to note that these were simply predicted task times, which calls their validity into question. Because of this, I am hesitant to make any high-level interpretation of the current results, other than to say that these problems could be addressed more directly via the inclusion of a pre-test phase in which participants would complete all possible task types so that their task times could be recorded.

CHAPTER 4. GENERAL DISCUSSION

Though we are able to choose between cognitive and physical costs, little is known about the relation between them. The experiments just presented are the first time, as far as I know, that observable physical actions have been used to compare physical and cognitive costs. This method is especially promising when one considers the sophisticated movement-tracking systems that are currently available. Although movement costs are rather broadly defined in the current study, the use of a movement-tracking system would allow for a more granular measure of physical variables, and therefore a more refined measure of cognitive cost. Further development of this method could be especially useful to clinical populations as a way to quantify cognitive deficits other than comparing an individual's performance in a task to mean performance.

A longer term goal of the present series of experiments is to shed light on the common substrate that allows us to flexibly plan between domains that share very little in common. When comparing one physical task to another, one has a myriad of common variables to draw from, such as the physical distance one must traverse during a movement, the muscular tension one might feel, how quickly one expects to become fatigued, and so forth. When comparing a physical task to a cognitive task, or when comparing across modalities more generally, the shared factors are often far less plentiful. This raises the question of what factors are shared across different modalities. Drawing upon the current results, it seems that one possibility is that such multi-modal comparisons are made using the difference in measured time between the two tasks. This is sensible, as cognitive and physical always share the factor of time in common.

However, there is a wrinkle in the time-based hypothesis of effort-estimation. A purely time-based strategy does not take into account very complex or difficult tasks that take only a short time, or vice versa. To recall the prior example, imagine that the comparison is between lifting 100 lbs. of weight and counting aloud to 30. Lifting 100 lbs. would likely take less time than counting to 30. Therefore, if measured time is the psychological index of effort, then lifting 100 lbs. of weight should feel subjectively easier than counting 30 digits aloud. This example raises two potential weaknesses of a purely-time based account of multi-modal effort estimation. The first is that, when considering rather extreme examples like the one just described, the predictions drawn from a time-based account do not seem to map on to predicted subjective effort. It is difficult to imagine that performing such a strenuous lift would, in fact, feel easier than rattling off a series of digits at a leisurely pace. The second problem is that a time-based strategy seems to predict sub-optimal behavior in terms of energy expense. The physical activity in the hypothetical scenario just described (lifting 100 lbs. of weight) could involve much higher weight values, as long as the lift takes less time than the cognitive task to which it is being compared. To push the example to the extreme: Would a person capable of lifting 500 lbs. in a short period of time really prefer to do so rather than count a series of digits? For obvious reasons, making metabolically costly decisions of this kind would be sub-optimal.

If participants do tend to switch to a different form of effort-estimation as task demands increase, as the time-based analysis of Experiments 4-6 subtly suggests, then what might this emergent strategy be? Extensive conversations with members of Penn State's Lab for Cognition and Action have led me to consider three additional possibilities for how one might compare cognitive and physical costs. The first possibility is that what matters is not *measured* time, as was tested here, but rather *subjective* time. Transitioning from measured to psychological time carries some benefits. It is possible that, much like the idea that expected effort can affect the perception of distance (Proffitt et al., 2003) so too can felt or predicted effort influence the perception of time. To return momentarily to the idea of ego depletion, Vohs and Schmeichel (2003) found that participants who regulated emotion while watching a video reported that the video felt longer than those who did not. These findings are consistent with the idea that the experience of time may expand and contract as a function of subjective effort. To return to the previous example, it is possible that lifting 100 lbs. could be objectively shorter in duration but could *feel* subjectively longer than counting aloud to 30.

It is important to note, however, that a subjective-time based model of effort estimation, as well as the results from Vohs and Schmeichel (2003) just described, clash with a popular model of time perception put forth by Church, Gibbon and Meck (1984). According to this model, the accumulation of pulse-based markers of psychological time underlies our ability to perceive time. The more pulses that accumulate, the longer the felt duration of time. Critically, tracking these internal markers of time requires some attention. If fewer attentional resources are available to track time, as would likely be the case during a tough task, fewer pulses can accumulate, leading to shorter estimated durations of time. Conversely, if more attentional resources are available to track time, more pulses could accumulate, leading to longer estimated durations. If subjective time underlies felt effort, this would predict that tasks requiring a lot of attention should feel not only shorter in terms of time, but also easier in terms of effort, while tasks that require little attention should feel longer in terms of time and more difficult in terms of effort. Therefore, at least according to the Church, Gibbon and Meck model of time perception, a subjective-time based strategy would seem to predict highly counterintuitive estimations of effort.

So how, then, do we compare different modes of effort? A promising possibility is that felt effort is the amount of *monitoring* one must do while a task is being completed (Rosenbaum, Potts & Muir, in prep). Such an account would nicely explain the aforementioned example concerning heavy weight loads and counting digits. Hefting 100 lbs. of weight is *hard*. The stakes involved with this task are high: there is a high probability of failing, and the consequences of failing are dire. A slip of the hand could mean a broken foot, an injured back, or worse. By comparison, the consequences of failing at counting aloud are relatively low; missing a digit might be embarrassing, but it does not mean a broken bone. One could argue, then, that to avoid the potential consequences associated with the task, hefting a heavy weight load requires a greater degree of monitoring than does counting aloud. Such a notion would accord with neuroscientific evidence suggesting that the anterior cingulate cortex is involved with both monitoring and effort perception (Mulert et al., 2005). Additionally, a monitoring-based hypothesis of effort estimation could bring some clarity to the current results. It is possible that the numerical decline in the mean *p*(Bucket) across weight conditions reflected a preference to avoid the added monitoring necessary to heft the heavy bucket to the end of the alley. Of course, such an interpretation is purely theoretical: the monitoring required to complete either task was not tested here.

Two methods come to mind that could address a monitoring-based account of effort estimation in future experiments. The first is simply to ask participants to rate the amount of monitoring that the tasks of interest required once they are finished. The second draws upon the logic that, if one is already heavily monitoring a task, one will be less able to monitor a concurrent task. For example, one could ask participants to wear a headband that would emit a number of subtle vibrations at random times during a primary task. The participant would be asked to track the number of vibrations, and to report that number either at a random time during or at the end of the primary task. The prediction would be that the more monitoring the primary task requires, the less able participants would be to successfully report the number of vibrations. Of course, it would be important to adjust the relevant cognitive tasks and physical tasks so that one is not "double-dipping" from the same general mechanism, so to speak: Counting digits aloud while also counting concurrent vibrations could be harder to coordinate than hefting buckets and counting vibrations. An especially promising feature of such a method is that one could, potentially, track non-linear features of monitoring across the course of a task if participants were interrupted and asked to report the current number of felt vibrations at random intervals throughout the primary task.

To sum up, the current results suggest that participants were accurate when they reported using time to compare between cognitive and physical tasks. Such a strategy is sensible: both cognitive and physical tasks take some time to complete. However, at some point, one must also factor in task demands when making decisions of this kind. If not, one runs the risk of falling into the trap of expending effort unnecessarily to save time. Thus, it is possible that, as the asymmetry in task demands increases, people rely not only on time, but on the amount of monitoring required to perform a task. A monitoring based account of effort is especially exciting when one takes into account current clinical knowledge. Might those who suffer from depression feel a sense of fatigue because they are so frequently engaged in a feature of self-directed attention: rumination (Takano & Tanno, 2009)? Along the same lines, might those who are diagnosed with ADHD experience heightened levels of energy because they are not expending as much effort via continued monitoring of a task? These questions and the theoretical implications that could follow highlight the importance of further tests of the monitoring account of multi-modal effort estimation.

REFERENCES

Backs, R. W. (1995). Going beyond heart rate: autonomic space and cardiovascular assessment of mental workload. *The international journal of aviation psychology*, *5*(1), 25-48.

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: is the active self a limited resource?. *Journal of personality and social psychology*, *74*(5), 1252.

Borg, G. A. (1962). *Physical performance and perceived exertion* (pp. 1-63). Lund: Gleerup.

Custers, R., & Aarts, H. (2005). Positive affect as implicit motivator: on the nonconscious operation of behavioral goals. *Journal of personality and social psychology, 89*(2), 129.

Davey, G. C., & Cleland, G. G. (1982). Topography of signal-centered behavior in the rat: Effects of deprivation state and reinforcer type. *Journal of the Experimental Analysis of Behavior*, *38*(3), 291.

Fiske, S. T., & Taylor, S. E. (1991). Social cognition, 2nd. *NY: McGraw-Hill*, 16-15.

Gailliot, M. T., & Baumeister, R. F. (2007). The physiology of willpower: Linking blood glucose to self-control. *Personality and Social Psychology Review*, *11*(4), 303-327.

Gailliot, M. T., Baumeister, R. F., DeWall, C. N., Maner, J. K., Plant, E. A., Tice, D. M., & Schmeichel, B. J. (2007). Self-control relies on glucose as a limited energy source willpower is more than a metaphor. *Journal of personality and social psychology*, *92*(2), 325.

Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. *Annals of the New York Academy of sciences*, *423*(1), 52-77.

Gray, W. D., & Boehm-Davis, D. A. (2000). Milliseconds matter: An introduction to microstrategies and to their use in describing and predicting interactive behavior. *Journal of Experimental Psychology: Applied*, *6*(4), 322.

Guidetti, L., Sgadari, A., Buzzachera, C. F., Broccatelli, M., Utter, A. C., Goss, F. L., & Baldari, C. (2011). Validation of the OMNI-Cycle Scale of Perceived Exertion in the elderly. *Journal of aging and physical activity*, *19*(3), 214-224.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology*, *52*, 139-183.

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive science*, *18*(4), 513-549.

Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, *36*(06), 661-679.

Macrae, C. N., Milne, A. B., & Bodenhausen, G. V. (1994). Stereotypes as energy-saving devices: a peek inside the cognitive toolbox. *Journal of Personality and Social Psychology*, *66*(1), 37.

Marshall, S. P. (2002). The Index of Cognitive activity: Measuring Cognitive workload. In *2002 IEEE conference on human factors and power plants* (pp. 7-5).

Mulert, C., Menzinger, E., Leicht, G., Pogarell, O., & Hegerl, U. (2005). Evidence for a close relationship between conscious effort and anterior cingulate cortex activity. *International Journal of Psychophysiology*, *56*(1), 65-80.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), 97-113.

Proffitt, D. R., Stefanucci, J., Banton, T., & Epstein, W. (2003). The role of effort in perceiving distance. *Psychological Science*, *14*(2), 106-112.

Reid, G. B., & Nygren, T. E. (1988). The subjective workload assessment technique: A scaling procedure for measuring mental workload. *Advances in psychology*, *52*, 185-218.

Robertson, R. J., & Noble, B. J. (1997). Perception of physical exertion: methods, mediators, and applications. *Exercise and sport sciences reviews*, *25*, 407.

Robinson, M. M., & Morsella, E. (2014). The subjective effort of everyday mental tasks: Attending, assessing, and choosing. *Motivation and Emotion*, *38*(6), 832-843.

Rosenbaum, D. A. (2008). Reaching while walking: Reaching distance costs more than walking distance. *Psychonomic Bulletin & Review*, *15*(6), 1100-1104.

Rosenbaum, D. A., Chapman, K. M., Weigelt, M., Weiss, D. J., & van der Wel, R. (2012). Cognition, action, and object manipulation. *Psychological bulletin*, *138*(5), 924.

Rosenbaum, D. A., Gong, L., & Potts, C. A. (2014). Pre-crastination: Hastening subgoal completion at the expense of extra physical effort. *Psychological Science*, 25, 1487-1496.

Rosenbaum, D. A. & Gregory, R. W. (2002). Development of a method for measuring moving-related effort: Biomechanical considerations and implications for Fitts' Law. *Experimental Brain Research*, 142, 365-373.

Rosenbaum, D.A., Potts, C.A. & Muir, M.A. (2016). The apples-and-oranges problem for task comparison: Why is it so easy to compare the difficulty of different kinds of tasks and to decide which task should be done first? Manuscript in preparation.

Takano, K., & Tanno, Y. (2009). Self-rumination, self-reflection, and depression: Self rumination counteracts the adaptive effect of self-reflection. *Behaviour Research and Therapy*, *47*(3), 260-264.

Tice, D. M., Baumeister, R. F., Shmueli, D., & Muraven, M. (2007). Restoring the self: Positive affect helps improve self-regulation following ego depletion. *Journal of Experimental Social Psychology*, *43*(3), 379-384.

Tomie, A., Grimes, K. L., & Pohorecky, L. A. (2008). Behavioral characteristics and neurobiological substrates shared by Pavlovian sign-tracking and drug abuse. *Brain research reviews*, *58*(1), 121-135.

Vohs, K. D., & Schmeichel, B. J. (2003). Self-regulation and extended now: Controlling the self alters the subjective experience of time. *Journal of personality and social psychology*, *85*(2), 217.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic bulletin & review*, *9*(4), 625-636.