

How are cognitive and physical difficulty compared?

Cory Adam Potts¹ · Stefan Pastel² · David A. Rosenbaum^{1,3}

Published online: 7 November 2017
© The Psychonomic Society, Inc. 2017

Abstract Tasks that require less physical effort are generally preferred over more physically demanding alternatives. Similarly, tasks that require less mental effort are generally preferred over more mentally demanding alternatives. But what happens when one must choose between tasks that entail different kinds of effort, one mainly physical (e.g., carrying buckets) and the other mainly mental (e.g., counting)? We asked participants to choose between a bucket-carrying task and a counting task. Our participants were less likely to choose the bucket task when it required a long reach rather than a short reach, and our participants were also less likely to choose the bucket task the smaller the final count value. We tested the hypothesis that subjective task durations provided a common currency for comparing the difficulties of the two kinds of tasks. We found that this hypothesis provided a better account of the task choice data than did an account that relied on objective task durations. Our study opens the door to a new problem in the study of attention, perception, and psychophysics—judging the difficulty of different kinds of tasks. The approach we took to this problem, which relies on two-alternative forced choice along with modeling the basis for the choice, may prove useful in future investigations.

Keywords Counting · Decision-making · Effort · Motor control · Perception and action · Reaching · Walking

If you were asked to choose the easier of two mainly physical tasks—reaching near or reaching far for a bucket to be carried some distance—you would probably pick the near reach. Similarly, if you were asked to choose the easier of two mainly mental tasks—counting up to 8 or counting up to 20, you would probably count to 8. But what if you were asked to choose between one of the bucket tasks and one of the counting tasks? If you could count up to 8 or reach far to pick up and carry a bucket, you would probably pick the counting task. But if you could count up to 20 or reach *near* to pick up and carry a bucket, you might be more likely to choose the bucket task. Regardless of whether you would behave as just suggested, it is likely that you would be able to make the choices relatively easily. How would you do so?

Surprisingly, there has been very little research on multi-modal task difficulty. This is surprising given the enduring interest in the factors contributing to the difficulty of physical tasks (e.g., Fitts, 1954; Rosenbaum, 2012), the enduring interest in the factors contributing to the difficulty of mental tasks (e.g., Kool, McGuire, Rosen, & Botvinick, 2010), and the longstanding interest in, and current explosion of interest in, multisensory integration (e.g., Rosenblum, 2010; Stevens, 1961). Given the growing appreciation that the divide between intellectual activity and physical activity may be illusory (Barsalou, 2008; Eisenberger, 1992; Rosenbaum, Carlson, & Gilmore, 2001; Rosenbaum, 2017; Schmidt & Bjork, 1992), it is fitting that attention be paid to the assessment of task difficulty for tasks that load more heavily on physical capabilities or on mental capabilities. In reality, both kinds of tasks draw on both kinds of capabilities. Even simple physical actions turn out to be far more cognitively sophisticated than might be supposed (Rosenbaum, 2017). In addition, cognitive tasks generally require some physical action, such as pressing one button or another or vocalizing a choice. From this perspective, it would be unsurprising if people can readily decide which of two tasks

✉ Cory Adam Potts
corya.potts@gmail.com

¹ Pennsylvania State University, University Park, PA 16802, USA

² Bielefeld University, Bielefeld, Germany

³ University of California, Riverside, CA 92521, USA

is harder or easier, even if, superficially, the two tasks seem to be of very different kinds.

Because so little attention has been paid to the nature of multimodal task difficulty, we began by considering possible ways that multimodal tasks might be compared. Our main question was, What is the common currency for comparing physical-task difficulty and intellectual-task difficulty? We assumed that there must be some currency, and we had a simple reason for supposing so: Common units of measurement must be used to compare quantities, as every student learns in elementary physics classes. Centimeters and inches cannot be added, for example; there must be conversion of one unit to the other. Similarly, adding centimeters and liters is unthinkable unless there is some special means of relating them (e.g., how high the surface of a bottled liquid is depending on its volume). Even if the quantities being compared are dimensionless (i.e., they have no units of measurement because they are divided by a reference value, allowing the units of measure to cancel out), the dimensionless values from different sources must still be related by some common currency to allow the values to be weighted according to their relative costs or benefits.

It is tempting to suppose that there is one common currency for judging all tasks, but we are skeptical that there is such a common currency or that, if it exists, it can ever be found. The common currency for judging the relative difficulty of different kinds of tasks might vary, and different individuals might deal with the same task comparisons in different ways. For example, if time were at a premium in one context, people might use task-completion times as the basis for choosing between the tasks, but if burning calories were what mattered most, the same people might make energy consumption the relevant basis for deciding what to do.

These remarks need not suggest that looking for common currencies is a fool's errand. Presumably, regularities can be found in the way people assess multimodal task difficulty. Most or even all people might tend to use the same common currency in a given context.

We thought it was worth pursuing this possibility in order to draw attention to the general problem of multimodal task difficulty and to show how, methodologically, one can investigate this topic. The method we used is the familiar two-alternative forced choice (2AFC) procedure. By varying the nature of two tasks between which participants choose, we sought to infer the relative importance of the possible bases for the judgment. Our lab has used this method before to study choices between physical tasks, including ones using different action modalities—namely, *reaching* over some distance versus *walking* over some distance (Rosenbaum, 2008; Rosenbaum, Brach, & Semenov, 2011; Rosenbaum, Gong, & Potts, 2014). But we have never extended the method to choosing between tasks

that are ostensibly of entirely different kinds—one “more mental” and one “more physical.”¹

What, then, might be the common currency for judging the relative difficulty of mental and physical tasks? One candidate is probability of success. Tasks might be defined as easy because they have high probabilities of success or difficult because they have low probabilities of success. This seems like a straightforward, reasonable possibility, but on reflection we came to doubt its usefulness. Probability of success is unlikely to be the sine qua non of task difficulty because tasks that differ dramatically in subjective difficulty can have the same probabilities of success. Think of rolling a huge boulder up a hill versus pitching a penny into a remote tiny hole. The probability of success for the two tasks might be the same, but boulder rolling is surely harder than penny pitching.

Another problem with probability of success is that defining success or failure may be elusive; estimating probability of success may be difficult or even impossible. Walking a long way might have a high probability of success in that it is virtually certain one will reach one's destination, but the subjective difficulty of a very long walk is clearly different from the subjective difficulty of a very short walk whose chance of completion is just as high.

A second candidate for the common currency of physical and mental task difficulty is the expected value of the task (i.e., the probability of success multiplied by its valence). This alternative account may help explain why rolling a boulder up a hill seems harder than pitching a penny to a far-away tiny hole, though the two activities are equally likely to lead to success or failure. The valence of boulder rolling is more negative than the valence of penny pitching. In common parlance, boulder rolling is harder than penny pitching. But saying this begs the question of where the valences come from. It gets one nowhere to say that boulder rolling has stronger negative valence than penny pitching because, say, one needs to be paid more to get people to roll boulders than to toss coins (cf. Westbrook & Braver, 2015; Westbrook, Kester & Braver, 2013). Saying that one must be paid more begs the question of why.

A third possibility is that the smaller the number of survival-related options available during a task, the harder the task is judged to be (Kurzban, Duckworth, Kable, & Myers, 2013). We find much to like about this hypothesis,

¹ We became interested in this topic when our lab discovered *pre-crastination*, the tendency to hasten task completion even at the expense of extra physical effort (Rosenbaum et al., 2014). In those experiments, participants chose between picking up and carrying one of two buckets. Surprisingly, most participants chose the bucket nearer to them at their start position, meaning it had to be carried farther than the other bucket. Based on subjects' comments and a series of tests designed to rule out other explanations of this result, Rosenbaum et al. suggested that participants traded physical work for cognitive work. Participants apparently were willing to take on more physical load to reduce the load on working memory of having to remember to pick up a bucket. The experiments reported in the present article were not designed to test this explanation of pre-crastination, though they were triggered by the desire to better understand physical versus mental trade-offs.

for it provides a straightforward explanation of the greater subjective difficulty of boulder rolling than penny pitching (among other contrasts). The energy required for boulder rolling is greater than it is for penny pitching, so one might be less equipped to handle other survival-related challenges when pushing a boulder to a summit than when pitching pennies at an arcade. Still, a challenge for the opportunity-cost model is how to count the number of survival-related options.

A fourth possibility, and the one we took most seriously, is that the effort of a task depends on the time spent on it. This proposal builds on an earlier suggestion by Gray and Fu (2004) and Gray, Sims, Fu, and Schoelles (2006) that cognitive strategies are selected more often than perceptual-motor strategies when the cognitive strategies take less time. To the best of our knowledge, Gray and colleagues' argument is one of the very few that has specifically addressed the problem of cross-modal task difficulty (but see Ballard, Hayhoe, & Pelz, 1995; Ballard, Hayhoe, Pook, & Rao, 1997; Wilson, 2002).

We find the time hypothesis promising for two main reasons. First, time is amodal and so, in principle, provides a natural bridge between ostensibly incommensurate quantities, such as mental and physical effort. For example, in connection with the comparison of centimeters and liters, one might link the two by referring to the time needed to raise the surface of a liquid being poured into a bottle. Second, the experience of time is psychologically rich. Subjective time does not equal objective time, but instead depends on a host of factors, including the amount of attention given to the events whose durations are judged (Block & Gruber, 2014; Zakay & Block, 1996). It is known that time alone does not dictate task difficulty (Kool et al., 2010). Instead, time might provide an effort-related cue for a more global metacognitive evaluation of task difficulty (Dunn, Lutes, & Risko, 2016; Dunn & Risko, 2016). In accord with the view of Dunn and colleagues, who have suggested that perceived demand serves as a metacognitive "summary" variable that indexes performance-related variables, we hypothesize that subjective duration might share a similar function. Specifically, we predict that participants' judgments of duration will expand and contract as a function of other performance-related variables, such as time and physical demand. All of these considerations encourage evaluation of the hypothesis that the subjective durations of tasks might provide a basis for judging the tasks' difficulty.

Experiment 1

In the first experiment, we asked subjects to choose between two tasks in each trial. One task was, from an intuitive standpoint, mainly physical, and the other was, also from an intuitive standpoint, mainly mental. We asked our subjects to choose between picking up and carrying a bucket and

counting up to various target values. The factors distinguishing the bucket tasks were how far subjects had to reach to get the bucket, how heavy the bucket was, and which hand had to be used to pick up the bucket and carry it. We chose bucket carrying as our mainly physical task because of previous work from our lab on the subjective difficulty and coordination of reaching and walking (e.g., Rosenbaum, 2012). Because walking and reaching are common, well-coordinated actions (van der Wel & Rosenbaum, 2007), we sought a mainly cognitive task that we thought, purely intuitively, would be of roughly comparable difficulty. Roughly equating the difficulty of the tasks of different kinds despite the absence of previous, formal data on the subject, let us try to avoid ceiling (or floor) effects in our choice probabilities. We chose counting as our cognitive task, owing in part to its automaticity (Naparstek & Henik, 2010), but also on the basis that, again, intuitively, we thought the chance of failure on the counting task would be comparable (equally low) as the chance of failure on the bucket task. We recorded the probability of choosing each version of the bucket task (one minus the probability of choosing the counting task) when it was paired with each version of the counting task (counting up to 8, 12, 16, or 20). We also recorded the times to complete the chosen tasks.

Method

The general setup appears in Fig. 1. We asked our participants either to walk at a leisurely pace, pick up a bucket from a table, and carry it to the end of an alley, or count aloud by ones from 1 at a leisurely pace up to a target value of 8, 12, 16, or 20. Besides varying the difficulty of the counting task, we varied factors that we thought might affect the difficulty of the bucket task. One was the side the bucket—either on the left or on the right. We asked participants to pick up and carry the bucket on the left side (if that is where the bucket was) with the left hand, or to pick up the bucket on the right side (if that is where the bucket was) with the right hand. Because most of our participants were right-handed, we thought right-side pickups would be easier than left-side pickups. Therefore, we expected that, all else being equal, our participants would be more likely to pick up the right bucket than the left bucket when the alternative task was counting.

We also varied the load in the bucket. For one group of participants, the bucket was empty. For another group, the bucket had 3.5 pounds of pennies. For a third group, the bucket had 7.0 pounds of pennies. We thought our participants would be more likely to pick the counting task when they had to lift a heavy bucket than when they had to lift a light or empty bucket.

A third physical-task factor we varied was the distance of the bucket from the edge of the alley. The bucket was either adjacent to the edge of the alley (.15 m), so in easy reach, or

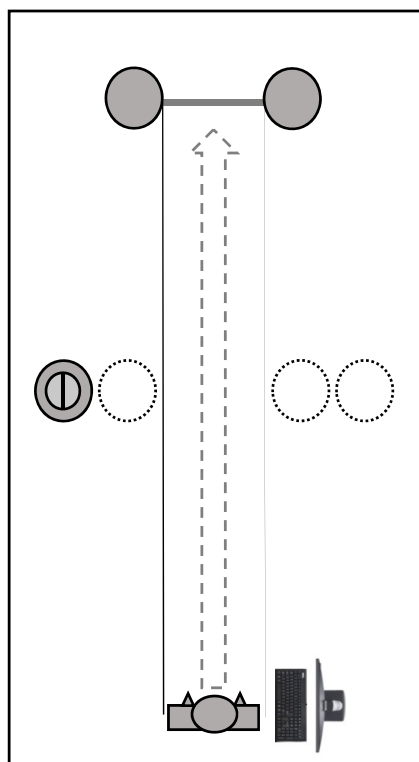


Fig. 1 Schematic of the experimental setup. The participant (figure at bottom) used a keyboard to select whether he or she would carry the bucket (shown here at only one of four possible locations) to its end table, or count aloud to 8, 12, 16, or 20. Just one target value was available per trial. Choices were displayed on the computer monitor, shown here on the right (beside the keyboard), though it and the keyboard were on the left for a random half of the subjects

far from the edge of the alley (.71 m), so requiring a long reach because subjects had to lean over a string boundary (shown as black lines on either side of the alley in Fig. 1) to get the bucket. Previous work in our lab (Rosenbaum, 2008, 2014; Rosenbaum et al., 2011) has shown that people strongly prefer short reaches to long reaches when they have to pick up an object (actually, the same child's beach bucket as used here). As shown in the previous work, people walked far to pick up the bucket with a short reach, favoring that action over walking a short distance to pick up the same bucket with a long reach. Accordingly, we expected that in the present experiment participants would treat long-reach tasks as harder than short-reach tasks, so they would be more likely to prefer counting to reaching when the reach was long than when the reach was short.

Participants

We tested 24 participants in each of the three physical load conditions, for a total of 72 participants (54 female, 18 male, mean age = 19.43 years, range: 18–34). We chose this sample size in accord with previous two-alternative forced-choice (2AFC) studies from our lab. Sixty-six of the 72 participants

tested were right-handed, as determined by a score of 7 or higher on the Edinburgh handedness inventory (Oldfield, 1971). Two participants in each of the three physical load groups were left-handed, as determined by a score of 7 or lower. The study was approved by the Penn State Institutional Review Board. Participants in Experiment 1 as well as Experiment 2 were compensated with course credit for their participation.

Apparatus and procedure

The participant stood at one end of a 2-foot (0.61 m) wide alley bordered to the left and right by white cotton string. One end of each of the two strings was attached to a 24-in. (0.12 m) high vertically placed .75-in. (1.91 cm) diameter plastic pipe that functioned as a post at the start of the alley. The other end was attached to the inside edge of the table standing on the same side at the end of the alley. A single bucket occupied a table 8 feet (2.44 m) from the participant's start position. The bucket stood on the left or right side of the alley and was either adjacent to the edge of the alley (.15 m) or at roughly 80% of the participant's average arm length (mean = .71 m) away from the alley's near edge. Pilot participants' arm lengths were measured from the acromion (the bony protrusion extending over the shoulder joint) to the tip of the middle finger. The bucket was a bright yellow plastic beach pail, 5-in. (12.7 cm) high, with a base 4 in. (10 cm) in diameter and a top 7 in. (17.8 cm) in diameter. The circular table on which the bucket stood was 24-in. (0.12 m) high and 36 in. (0.91 m) in diameter. The two additional far tables at the end of the alley stood 16 ft. (4.88 m) from the start position. The far table that stood on the same side as the bucket in a given trial was the target table for that trial. In each trial, the bucket's upright dark-blue handle stood perpendicular to the long edge of the alley. All participants lifted the bucket to feel its weight before the experimental session began. Depending on which group the participant was in (based on random assignment), the bucket contained no added weight, 3.5 lbs (1.59 kg) of pennies, or 7.0 lbs (3.18 kg) of pennies. The inside of the bucket was occluded by a blue foam lid to hide the load. This was done to encourage the participants to make their choices based on the felt rather than the seen load. The blue foam load was on the bucket even when the bucket was empty.

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, 12 inches (30.48 cm) to the left for half the participants or the same distance to the right for the other participants. In each trial, a choice appeared on this computer monitor. Each choice consisted of one physical task, namely, carrying the bucket from its current position to its corresponding end table, and one cognitive task, namely, counting aloud by ones to 8, 12, 16, or 20. Drawing upon pilot work, these count values were

chosen such that, assuming participants would count at comparable rates, counting aloud to 8 and 12 should have taken less time, on average, than carrying the bucket, while counting aloud to 16 or 20 should have taken more time, on average, than carrying the bucket. Participants chose the rate at which they performed either task. The only instructions they received regarding task rates were to walk in a natural pace without stopping and to count clearly and evenly.

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, the 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, he or she said “Open,” and the participant opened his or her eyes. Once the participant opened his or her eyes, he or she pressed the Enter key to advance to the next screen, which displayed the pair of physical and cognitive tasks for that trial. An example of a typical choice was, “Would you rather carry the bucket to the far table (press the *b* key) or count aloud to 12 (press the *c* key)?” The participant made his or her choice by pressing either the *b* key for “bucket,” or the *c* key for “count,” followed by the Enter key, and immediately began the task. This press of the Enter key started a timer programmed in MATLAB, which was not visible to the participant. Once the participant finished the chosen task, the participant pressed the Enter key once more, which stopped the timer. Then, the participant closed his or her eyes again to await the next trial. All possible combinations of bucket positions and count values were tested, resulting in a total of 16 trials (4 bucket positions \times 4 count values).

Results

Figure 2 shows the probability, $p(\text{Bucket})$, of performing the bucket task as a function of the count target for long reaches and short reaches, averaged over bucket side and bucket load. We averaged over bucket side and bucket load because a mixed-model ANOVA with three within-subject factors, Hand (left, right) \times Reach (short, long) \times Count (8, 12, 16, 20 digits) and one between-subjects factor (added weight: 0 lbs, 3.5 lbs, 7 lbs) showed no main effects or interactions involving bucket side or bucket load (all p s $>$.05), but highly significant main effects of reach distance, $F(1, 69) = 53.70$, $p < .001$, $\eta_p^2 = .438$, and count value, $F(2.33, 160.96) = 56.09$, $p < .001$, $\eta_p^2 = .630$. There was no interaction between these factors. As seen in Fig. 2, participants were less likely to choose to carry the bucket when it required a long reach ($M = .45$) than when it required a short reach ($M = .70$), and participants were less likely to

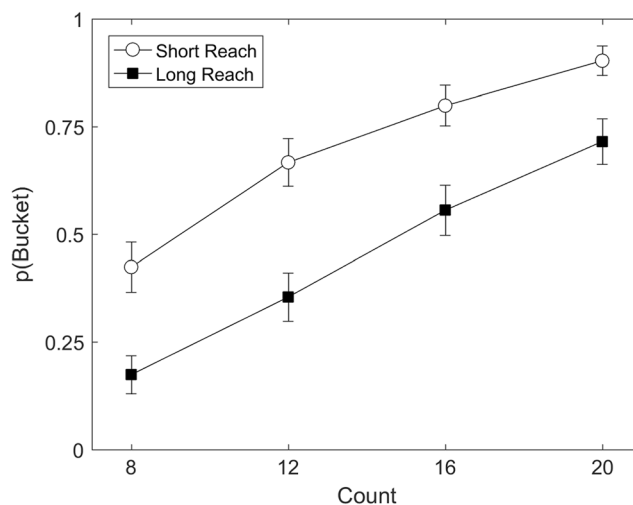


Fig. 2 Probability (± 1 SE) of reaching for and carrying the bucket, $p(\text{Bucket})$, plotted as a function of the number to be reached in counting (Count), averaged over hand (left or right) and added weight (0 lbs, 3.5 lbs, or 7.0 lbs). Data from Experiment 1

carry the bucket the shorter the count value: 8 digits ($M = .30$), 12 digits ($M = .51$), 16 digits ($M = .68$), and 20 digits ($M = .81$).

We also analyzed the times to complete the tasks. Recall that immediately before and after performing the chosen task the participant pressed the Enter key. We used the times between the two presses on the Enter key to provide an estimate of the time to complete the task in that trial. Figure 3 shows the mean time in seconds (s) for participants to perform the count task for each of the four possible target count values (left panel), and also to perform the bucket task for the two possible reaching distances (right panel). An ANOVA yielded a significant main effect of count value such that the counting task took longer as the count value increased, $F(3, 69) = 21.07$, $p < .001$, $\eta_p^2 = .478$. The time to complete the physical task was somewhat shorter for the short-reach tasks than for the long-reach tasks, but this difference was not statistically significant, $F(1, 45) = .366$, $p > .05$, $\eta_p^2 = .008$. We did not include the counting task and the bucket task in the same ANOVA because we had no specific hypothesis about the relation between the two kinds of task times.

Figure 4 shows the relation between $p(\text{Bucket})$ and task times. The measure of task times used was, for each task choice, the ratio of counting time to the sum of counting time and bucket time. Our use of this formulation instantiates the Luce Choice axiom (Luce, 1959), which our lab has used in all previous research we have done on task choices. For a review of our use of this measure, see Rosenbaum, Chapman, Coelho, Gong, and Studenka (2013).

The left panel of Fig. 4 shows $p(\text{Bucket})$ as a function of the ratio just referred to, using the times we recorded. As seen in the figure, $p(\text{Bucket})$ increased as the ratio increased, consistent with the view that as the count time increased relative

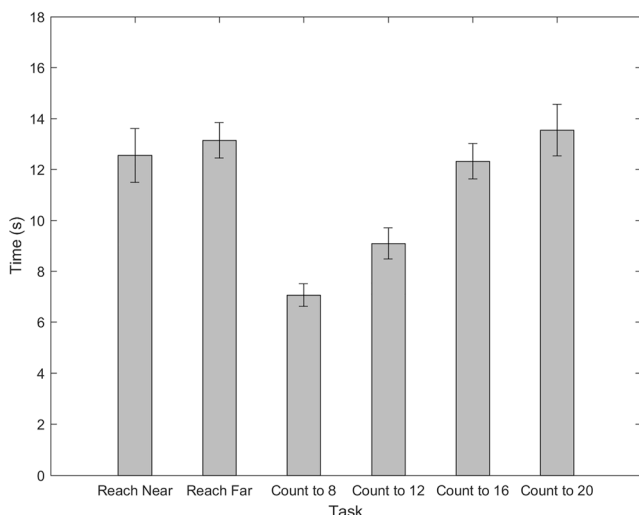


Fig. 3 Mean times (± 1 SE) to complete tasks with near and far reaches and count values up to 8, 12, 16, or 20. Data from Experiment 1

to the reach time, the probability of selecting the bucket task increased. The Pearson product-moment correlations between the ratios and the $p(\text{Bucket})$ values were $r = .97$ for the short reach condition and $r = .90$ for the long-reach condition, with $p < .001$ in both cases, but with $df = 21$ rather than $df = 23$ because two subjects always picked one task, meaning it was impossible to compute a task time difference for them.

We plotted the near-reach and far-reach conditions separately to bring out an important finding: The two curves were

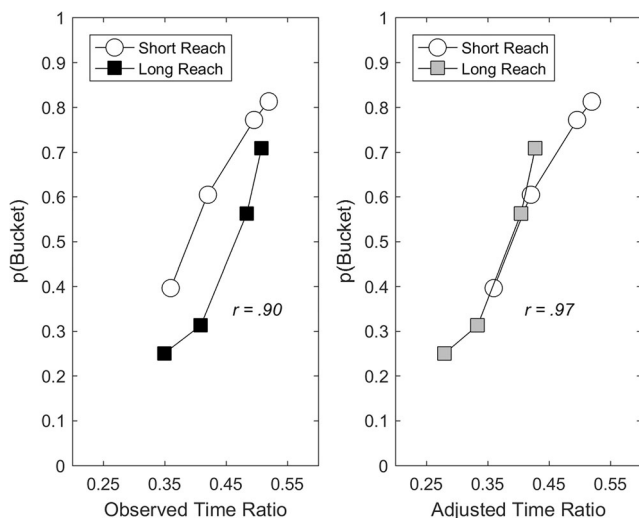


Fig. 4 Probability, $p(\text{Bucket})$, of choosing the bucket task as a function of the observed time ratio (left panel) and as a function of the adjusted time ratio (right panel). The observed time ratio is the mean time to complete the counting task (counting up to 8, 12, 16, or 20) divided by the sum of that time and the time to complete the short or long bucket-reaching task. The adjusted time ratio is the same as the observed time ratio but includes an extra hypothetical 5 seconds for the long reach task. Correlations are based on pooling the points within each panel. Actual data (the black squares in the left panels and the empty circles in both panels) are from Experiment 1. The gray squares in the right panel are hypothetical

separated, implying that actual performance time, by itself, did not predict task choice. Pooling the near-reach and far-reach data points yielded a correlation between the ratios and the observed values of $p(\text{Bucket})$, $r = .90$.

Discussion

In Experiment 1, we asked our subjects to choose between a mainly physical task (walk, pick up, and carry a bucket to the end of an alley) and a mainly cognitive task (count up to a target value). We found that our subjects preferred the bucket task to an increasing degree as the count target grew. As expected, it took longer for our participants to complete the counting task the larger the final count value. However, our participants’ task completion times were not significantly longer when long reaches were required than when close reaches were required. This outcome can be explained in terms of a well-established principle of visually guided aiming. When targets are farther away, the added time to reach the targets depends on the required aiming precision. The greater the required aiming precision, the greater the contribution of target distance to movement time (Fitts, 1954). In the present experiment, the required aiming precision was small compared to the high precision required in typical visual aiming tasks, where subjects try to move a pointer as quickly as possible to a small (often tiny) target on a screen. In the present experiment, the aiming requirement, such as it was, involved grabbing the standing handle of a beach bucket. Moreover, no instruction was given about speed. Therefore, it was understandable that whether the bucket was near or far did not matter to a statistically significant degree in terms of task completion time, though there was a tendency for the long-reach tasks to take somewhat longer than the short-reach tasks (an outcome that will be replicated and amplified a bit in the next experiment).

It may seem puzzling that while our participants chose to pick up the bucket less often when it required a longer reach, our participants’ choices did not differ across the three physical loads tested, nor between the left and right sides. These results raise the question of whether our participants were actually sensitive to physical costs. We think they were in view of results obtained by Rosenbaum et al. (2014). In that study, which was referred to earlier, it was found that participants preferred to carry a lighter bucket rather than a heavier bucket when they could choose between a bucket on the left or a bucket on the right and the buckets had unequal weights, though when the buckets were equally weighted and were equidistant from the start and end lines, participants preferred the right bucket, in accord with their being right-handed. Therefore, in situations like the one studied here, subjects from the general population we studied—Penn State University students in both cases—were in fact sensitive to load and to side (or hand). We speculate that the added load

used here was not heavy enough to significantly tax our subjects' physical capabilities. Presumably, if the bucket had weighed much more (e.g., 200 pounds rather than 7) and the alternative task was to count to 20, then no one, or virtually no one, would have chosen the bucket task.

Regarding the lack of difference between the left-hand and right-hand conditions, this result may be interpreted from the standpoint of the so-called dynamic-dominance hypothesis of handedness, proposed by Sainburg (2005). According to this hypothesis, the dominant hand is specialized for dynamics (e.g., bucket lifting) while the nondominant hand is specialized for statics (e.g., bucket holding). Both components were important here, so neither hand may have been favored.

Finally, what if anything can be said about the finding that the observed task performance times accounted for some but not all of the task choice data? Recall that $p(\text{Bucket})$ increased with the Luce ratios relating counting time to reaching time, but two curves emerged rather than one. Is there some way to bring the two curves together?

One possibility is that *objective* task times were not the basis for the task choices, but *subjective* task times were. Suppose, for example, that the subjective duration of the long-reach task was longer than the subjective duration of the short-reach task. A way to represent this outcome graphically is to focus on Fig. 4 and to shift the long-reach points to the left along the abscissa. We did this in the right panel of Fig. 4, effectively saying that the experienced duration of the long-reach task was amplified by 5 seconds. With this horizontal shift in the long-reach time-difference points, we could fit a single straight line to all the points, yielding a Pearson product-moment correlation of $r = .97$, $p < .001$, $df = 42$, up from $r = .90$. This difference in correlations is statistically significant, $z = -2.81$, $p < .01$, $df = 44$, (one-tailed). This outcome is consistent with the hypothesis that performance time, by itself, may have approximated the common currency used to pick the easier of the two tasks in Experiment 1, but a hypothetical transformation of the performance time, perhaps reflecting subjective time, could better predict participants' choices. In Experiment 2, we sought a more direct test of this hypothesis.

Experiment 2

In Experiment 2, we sought to replicate Experiment 1 and also to obtain subjective estimates of the times to do each task. Getting the subjective time estimates let us test the hypothesis that subjects judged long-reach tasks to take longer than short-reach tasks by an amount exceeding the actual time difference between the tasks. By obtaining subjective time estimates, we could also test the hypothesis that subjective times better predicted task choice probabilities than objective task times did.

Method

Thirty-six undergraduates from the University of California, Riverside (18 female, 18 male, mean age = 18.86 years, range: 18–22) performed the counting and bucket-carrying tasks from Experiment 1. Each participant performed in three contexts, with the order of the three contexts being balanced over subjects and with six subjects assigned at random to each of the six possible orders of the three contexts. One context was choosing between all the cognitive and physical task pairs (choice context), as in Experiment 1. Another context was estimating how long each task took (estimation context). A third context was performing the counting and bucket tasks without choosing between them (action context). Our aim in the action context was to obtain estimates of the times to carry out each task in a way that overcame a limitation of the procedure used in Experiment 1. There, the number of time estimates subjects gave was not the same for all tasks but depended on how often each task was chosen. We wanted to get an equal number of observations for each task to avoid possible self-selection artifacts (cf. Siegler & Lemaire, 1997). Thus, accounting for these three contexts, each participant performed a total of 48 trials in all (16 choice-context trials, 8 action-context trials, and 8 estimation-context trials, with the action and estimation context tasks each performed twice).

Experiment 2 was carried out at the University of California, Riverside, rather than at Penn State because the last author moved to UCR in the midst of this research. The setup at Penn State was recreated at UCR, although the space was quite different. Whereas the lab at Penn State was a normal indoor room, the space at UCR was a large outdoor, shaded arcade. The arrangement of materials was nearly identical, although no string boundary was present in Experiment 2. Instead, two tables, one closer to (.15 m) and one farther from (.71 m) the edge of the alley stood to the left and right side of the alley 8 feet (2.44 m) from the start position. Therefore, to pick up the bucket from the farther table, participants had to lean and reach over the table closer to the edge of the alley. Additionally, no computer was placed at the start of the subject's walkway in the UCR setup. Rather, the experimenter stood near the subject's start position and simply told the subject which task to do, or which tasks to choose between, in each trial. The subject's performance was digitally recorded (audio and visual) for off-line coding of his or her performance times, task choices, and time estimates. Performance times were defined as the duration between the time at which the participant's foot (the toe side of his or her shoe) visibly raised from the marked start position at the beginning of the trial and the time at which the participant's foot (toe) visibly crossed the start position on the way back.

There was one other change to the method. In Experiment 2, we eliminated the loaded-bucket conditions. Recall that bucket weight was not found to have a significant effect on

any dependent measure in Experiment 1. Bucket weight was also not found to have an effect in the choice experiments of Rosenbaum et al. (2014) when subjects chose between carrying one of two unloaded buckets or chose between carrying one of two loaded buckets with 3.5 pounds or 7 pounds of pennies each (as in the present Experiment 1). Given these earlier results, we decided to test just one group of subjects in Experiment 2. The one group tested in Experiment 2 used an unweighted bucket.

In terms of design and procedure, when subjects were asked to choose between the bucket task and the counting task, the bucket was set up without the subject watching, as in Experiment 1. While the experimenter prepared the position of the bucket for the next trial, the participant faced the direction opposite the experimental display until he or she received further instruction. The order of mention of the two possible tasks was balanced over trials. In the action and choice contexts, subjects were told to do the task as they normally would. In the estimation context, subjects were told to say how long they thought each task would take in seconds. The order of the tasks within each of the three contexts was random per subject. The experiment was approved by the UCR Institutional Review Board.

Results

Figure 5 shows $p(\text{Bucket})$ as a function of count value for the short-reach and long-reach conditions. The graph is restricted to these factors because no other factor had a significant effect on $p(\text{Bucket})$, as tested with a mixed-design ANOVA designed to evaluate the effects of the one between-subjects factor (six levels of task-type order) and three within-subjects factors: bucket side, reach distance, and count value. Bucket side (left or right) did not have a significant effect and

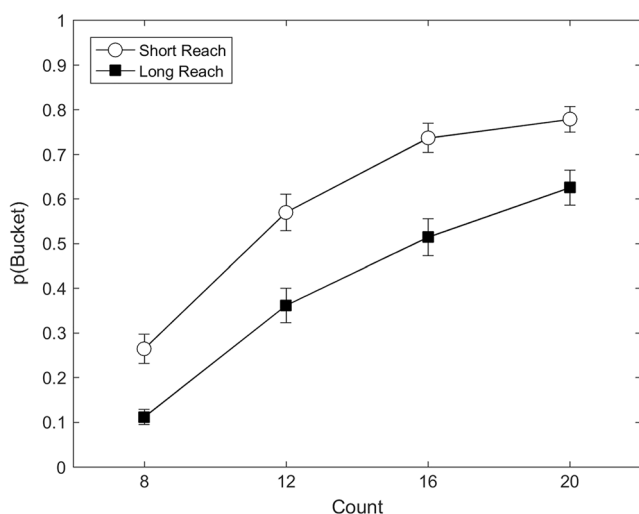


Fig. 5 Probability (± 1 SE) of reaching for and carrying the bucket, $p(\text{Bucket})$, plotted as a function of the number to be reached in counting (Count), averaged over hand (left or right). Data from Experiment 2

did not interact with any other factor. However, reach distance had a statistically significant effect, $F(1, 30) = 13.07, p < .01, \eta_p^2 = .303$, as did count value, $F(3, 64.61) = 29.46, p < .001, \eta_p^2 = .625$. There were no other main effects or interactions.

Figure 6 shows the objective and subjective times for the tasks. We analyzed these data in two separate ANOVAs, one for the bucket times and one for the count times, because we did not have any specific hypotheses about bucket times versus count times. We entered the objective bucket task times (i.e., the off-line, video-analyzed times between the first step from the start line to placement of the bucket on the relevant target at the end of the alley) and the subjective bucket task times (i.e., subjects' time estimates) into a mixed-design ANOVA designed to test the between-subjects effect of task-type order (six levels), and the within-subject effects of type of time (objective or subjective), bucket side (left or right), and bucket distance (near or far). The ANOVA yielded just two statistically significant results. There was a main effect of reach distance, $F(1, 29) = 55.48, p < .001, \eta_p^2 = .657$, with longer times for far reaches ($M = 10.67$) than for near reaches ($M = 9.19$), and a significant interaction between reach distance and objectivity/subjectivity of time, $F(1, 29) = 15.28, p < .01, \eta_p^2 = .345$. As seen in Fig. 6, times were longer for far reaches than for near reaches, and whereas subjective times were shorter than objective times for short reaches, subjective times were longer than objective times for long reaches.

Having just discussed the times to complete the bucket tasks, we turn now to the times to complete the count tasks. The count time data were analyzed in the analogous way, except that the within-subject factors were type of time (objective or actual, and subjective or estimated) and count value (four levels). The ANOVA yielded a main effect of type of time, $F(1, 29) = 11.17, p < .01, \eta_p^2 = .278$, with longer subjective durations ($M = 11.50$ s) than objective

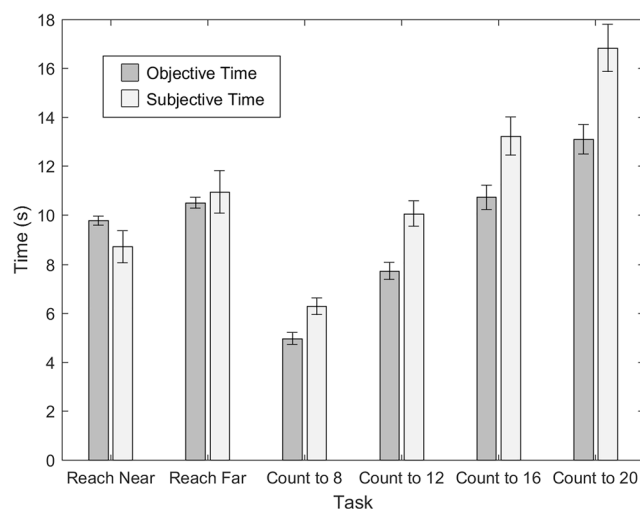


Fig. 6 Mean objective times (± 1 SE) and mean subjective times (± 1 SE) to complete tasks with near and far reaches and count up to 8, 12, 16, or 20. Data from Experiment 2.

durations ($M = 9.05$ s). Additionally, there was a significant main effect of count value, $F(3, 38) = 417.11$, $p < .001$, $\eta_p^2 = .935$, with longer actual and estimated time durations for higher count target values. Moreover, count value and time type interacted, $F(3, 87) = 5.72$, $p < .01$, $\eta_p^2 = .165$. As seen in Fig. 6, count time increased with count value, and subjective time exceeded objective time to a greater degree as the count value increased. The degrees of freedom reported are Greenhouse–Geisser corrected values to account for violations of the assumption of sphericity.

Figure 7 shows the relation between task completion times and task choice probabilities. As seen in the left panel of Fig. 7, choice probabilities were reasonably well predicted by the *objective* time ratios (i.e., the objective count times divided by the sum of objective count times and objective bucket task times). The Pearson product-moment correlation between the objective time ratios and $p(\text{Bucket})$ was $r = .93$. However, as seen in the right panel of Fig. 7, choice probabilities were better predicted by the *subjective* time ratios (i.e., the subjective count times divided by the sum of subjective count times and subjective bucket task times). The Pearson product-moment correlation between the subjective time ratios and $p(\text{Bucket})$ was $r = .98$. The difference between the two correlations was significant, $z = -1.7$, $p < .05$, based on a one-tailed test of the hypothesis that the subjective time correlation would be higher.

Discussion

Our aims in Experiment 2 were twofold. First, we wanted to see whether we could replicate the pattern of task choices observed in Experiment 1. We succeeded in that aim. As in Experiment 1, subjects elected to do the bucket task more often when the bucket task required a short reach than when it required a long reach. In addition, subjects chose the bucket task more the higher the target count value. The effects of reaching distance and count value were independent in both experiments, and there was no effect of the side of the bucket on the probability of choosing the bucket task in Experiment 1 and in Experiment 2.

Second, we wanted to test the hypothesis that the subjective durations of the tasks better predicted the choice data than the objective task durations did. We were led to this prediction by seeing two curves rather than one when we plotted $p(\text{Bucket})$ as a function of the time ratios in Fig. 4. The two curves corresponded to the near-reach and far-reach conditions and let us reject the hypothesis that objective time was the basis for choosing between the bucket task and counting task. We surmised that the two curves could be aligned if we allowed that the subjective durations of the long-reach task may have been longer than the subjective durations of the short-reach task, and that subjective durations rather than objective durations were used in subjects' implicit calculations. To test this

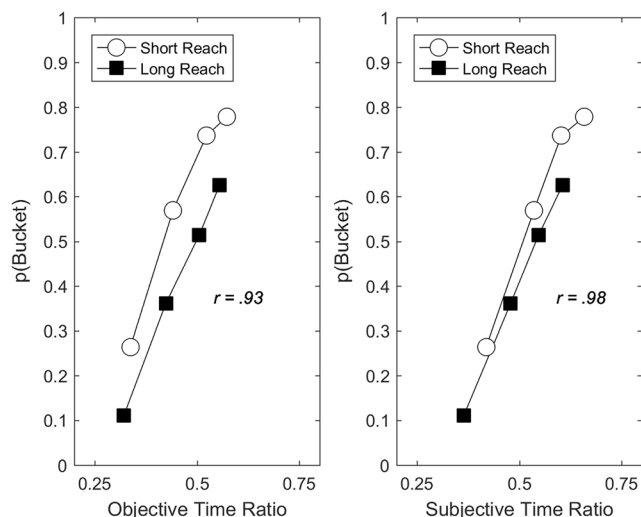


Fig. 7 Probability, $p(\text{Bucket})$, of choosing the bucket task as a function of the observed time ratio (left panel) and as a function of the subjective time ratio (right panel). The objective time ratio is the mean time to complete the counting task (counting up to 8, 12, 16, or 20) divided by the sum of that time and the time to complete the short-reach or the long-reach task. The subjective time ratio is the mean time estimate provided by the subjects to complete the counting task (counting up to 8, 12, 16, or 20) divided by the sum of that time and the mean time estimate provided by the subjects to complete the short-reach or long-reach task. All real data (no hypothetical data) from Experiment 2

hypothesis in Experiment 2, we collected subjective durations for each bucket and each counting task and measured objection durations for each task performed by the same subjects (based on the videos of their performance). Finally, we asked the same subjects to choose between each of the bucket tasks and each of the counting tasks.

The results of Experiment 2 were consistent with the hypothesis that subjective durations would better predict the task choices than would objective durations. As shown in Fig. 7, when we used *objective* times, we obtained *two* curves, but when we used *subjective* times, we obtained results more consistent with a *single* underlying curve, as demonstrated by a significantly higher correlation.

Was the latter outcome actually due to the far-reach task having a longer subjective duration than the near-reach task? It is worth considering the possibility that it was not because we used subjective durations for *all* the tasks when we generated the results in the right panel of Fig. 7 (counting tasks as well as bucket tasks), so it is not obvious that the far-reach subjective durations specifically accounted for the better fit. Several findings suggest that it was, however.

First, the far-reach task had a longer subjective duration than the near-reach task did, and this difference was larger than the accompanying objective time difference. Expanding on this point, it is worth noting that, for the *objective* times, the range of times for the far-reach tasks minus the near-reach tasks was small (10.5 s $-$ 9.8 s $=$ $.7$ s.), but for the *subjective* times, the range of times for the far reach tasks minus the near

reach tasks was much larger, more than three times as large ($10.9 \text{ s} - 8.7 = 2.2 \text{ s}$). The larger time difference better explains the robust difference in $p(\text{Bucket})$ for near as opposed to far reach tasks ($.59 - .40 = .19$).

This possibility is further supported by an added test in which we entered subjective bucket task durations along with objective count durations into the Luce ratio. In this case, we used a mixed model, one that combined subjective durations (for the bucket task) with objective durations (for the counting task). Our reasoning was that if the better fit obtained in the right panel of Fig. 7 compared to the left panel of Fig. 7 was preserved when the only subjective durations were for the bucket task, then that outcome would fit with the hypothesis that the subjective durations for the bucket tasks accounted for the improvement in fit. When we conducted this analysis, we found that the correlation between $p(\text{Bucket})$ and the time ratios remained at .98. Therefore, the bucket-time subjective durations were mainly responsible for the boost in correlation shown in the right side of Fig. 7.

It is worth noting, however, that, despite the evidence supporting the hypothesis that the subjective durations were exaggerated for the far-reach task, there is an aspect of our results that deviates from this hypothesis. As we predicted from Experiment 1, and as we noted previously, the subjective times for the far-reach task were longer than the objective times for the same, far-reach, task. However, as shown in the leftmost pair of bars in Fig. 6, participants also underestimated the near-reach task times. In fact, the absolute difference between the objective and subjective times was larger for the near-reach task (1.1 s) than for the far-reach task (.4 s). Said another way, the “time penalty” for the far-reach task that we predicted from the results of Experiment 1 did not emerge to the degree that we expected. However, recall that in Experiment 1 the only times that we had available were objective times, while our hypotheses related to subjective times. It is possible that the exaggerated time we predicted in Experiment 1 included both overestimations and underestimations of experienced time. What is most critical, in our view, is that the difference between the subjective times for the near-reach and far-reach tasks was larger than the difference between the objective times. One might speculate that the reason behind this enlarged difference is rooted, at least in part, in the relatively high psychophysical cost of reaching. Previous research comparing the costs of reaching and walking has suggested that reaching is approximately 11 times costlier than walking per unit distance (Rosenbaum, 2008). Because of the high cost associated with long reaches, participants might have reasoned that short reaches took less time than long reaches, highlighting the role of metacognitive evaluations in judgments of duration (cf. Dunn et al., 2016).

There is another possible interpretation of the role of subjective times in Experiment 2. This alternative interpretation rejects the direction of causation between subjective durations

and task choices that we have been assuming. We have implied that subjects’ subjective estimates of the durations of the tasks served as input to their decisions about which task to perform: The larger the subjective duration of one task relative to the other, the less likely the longer duration task was. The alternative interpretation is that the direction of causation was actually the other way around: Subjects recalled having preferred one task over another (for reasons other than time) and then inferred that the less preferred task took longer.

Our data do not fit with this account, however. If the account were correct, one would have expected an effect of task type order on subjective durations. That is, one would have expected the subjects’ estimates of the times to do the tasks to depend on whether they had already done them. We found no evidence for such an effect, however. The ANOVAs we performed on the times (both objective and subjective) and the ANOVA we performed on the task type order (act-choose-estimate, act-estimate-choose, choose-act-estimate, and so on) failed to turn up any effect of the ordering of task type (a between-subjects factor in our experimental design). From this outcome, we remain skeptical that our participants based their subjective durations on their task choices. Instead, we believe they based their task choices on their subjective experiences or anticipations of the tasks’ durations. Moreover, this result suggests that subjective judgments of duration did not differ when given in a prospective or retrospective manner.

General discussion

In everyday life, decisions are made all the time about which tasks to perform and when to perform them. The decisions often take into account the difficulty of the tasks. Insofar as the tasks have both physical and cognitive components, both components’ contributions are presumably taken into account in decision-making about them. Despite the commonness of this comparison process, and despite the deep theoretical interest that this problem holds, very little prior work has been done on it, as far as we know.

We sought to open an investigation of multimodal task-difficulty comparison by inviting participants to choose between a task that was “more cognitive” (counting) and a task that was “more physical” (picking up and carrying a bucket). We reasoned that if participants were sensitive to cognitive and physical costs, they would choose to perform the cognitive task more often as the physical task got harder, and similarly, they would choose to perform the physical task more often as the cognitive task got harder. We obtained data consistent with these predictions. Our participants selected the physical and cognitive tasks to different degrees depending on the tasks’ demands, along the lines just outlined. The common currency they apparently used were the tasks’ subjective durations.

We have three further points of discussion. The first relates to the tasks chosen and the common currency of subjective time. The second relates to the question of whether there was a general bias toward the more cognitive task or the more physical task in the present experiments. The third concerns whether participants' decisions reflected a decision to reduce subjective duration for a given trial or for the experiment as a whole.

With regard to the first topic, it is possible that our participants thought of the counting task as counting seconds in time rather than counting numerical values per se, in which case it would be unsurprising that a time-based metric applied to choices involving counting. We cannot conclusively rule out this possibility, but we think we have evidence against it. If participants were allotting roughly 1 second per digit, then one would expect the amount of time in seconds to complete the counting task to be comparable to the target count value. Said another way, counting to 8 should have taken about 8 seconds, counting to 12 should have taken about 12 seconds, and so on. This was not the case in either experiment. Count times were generally shorter for both the subjective and objective times than one would predict if participants had used this strategy. Moreover, and more importantly, subjective time exceeded objective time to a greater degree as the count value increased, suggesting that participants were not simply counting seconds and happened to have an imprecise notion of what one second is.

The second point of discussion concerns whether there was a general bias toward the more cognitive task or the more physical task in the present experiments. Some authors have suggested that there is a bias to conserve cognitive resources, sometimes at the expense of perceptual-motor resources (e.g., Wilson, 2002). Others (Gray et al., 2006) have argued against such a bias, claiming instead that if the time to complete either task is equal, no bias should appear. What is critical, according to these authors, is not the modality of the task but rather the cost incurred in terms of time. Referring to Fig. 4 and Fig. 7, in which $p(\text{Bucket})$ is plotted as a function of time ratios, the $p(\text{Bucket})$ value that corresponds to .5 of the abscissa (and so the point at which the bucket task and count task times were equal) is indicative of the presence or absence of such a bias. If the corresponding $p(\text{Bucket})$ value was above .5, that would suggest that, controlling for time, there was a bias to perform the bucket task. Values below .5 would suggest the same conclusion for the counting task. Although Fig. 4 suggests a "bucket bias" for Experiment 1—that is, $p(\text{Bucket})$ values corresponding to .5 of the abscissa were above .5, this finding did not emerge in Experiment 2. In fact, as shown in Fig. 7, the critical $p(\text{Bucket})$ values were near (left panel) or below (right panel) .5 on the y -axis. Therefore, because of the lack of a consistent bias across experiments, our results do not support the notion that there was a general preference to favor the more cognitive task over the more physical task or vice

versa. Rather, our results appear to accord more strongly with the time-based perspective proposed by Gray and colleagues.

The third point of discussion concerns the scale at which participants sought to minimize subjective task duration. In both experiments, participants could reduce the time spent on the experiment as a whole by reducing the time spent on each trial. Thus, although our results are consistent with the notion that participants preferred to minimize subjective duration on a trial-by-trial basis, it is unclear whether this strategy reflected a broader preference to reduce total time spent on the experiment. To tease apart these possibilities, one might increase the number of trials so that participants remain, or expect to remain, in the testing environment for the entire duration of the experimental session, regardless of the decisions they make in each trial. We do not, however, view a more global view of subjective time reduction as problematic to our perspective. Rather, such a strategy would underscore the potential generality of subjective duration as a currency for task selection across different domains and time scales.

A final comment is that although we obtained data consistent with the time-based account, we remain open to the possibility that some other variable was actually used—perhaps one or more of the variables discussed in the introduction of this manuscript. Subjective time may have been a proxy for another, more fundamental variable. The other variable we see as particularly interesting is attention. Attention is an important determinant of experienced duration (Block & Gruber, 2014; Zakay & Block, 1996), and attention and effort are closely linked (Kahneman, 1973). It has also been shown that cognitive effort is related to, but not identical to, cognitive task time (Kool et al., 2010). Finally, as mentioned before, Dunn and Risko (2016) and Dunn et al. (2016) have shown that time might provide an effort-related cue for metacognition. Insofar as attention and time are cross-modal, it is possible that attention serves as the basis for judging relative task difficulty. The data we have collected do not bear critically on this hypothesis. We hope to attend to this matter at some time in the future.

Acknowledgements We thank Bill Bui and Angus Cheng for help with data collection, and Richard Carlson, Christine Chiarello, John Franchak, Iman Feghhi, Karen Gasper, Jenna Monson, and Brad Wyble for helpful discussions. Frank Durgin, Chaz Firestone, and an anonymous reviewer gave valuable editorial feedback. We thank them especially for urging us to do the second experiment. The work was supported by a UCR CoR grant to the third author.

Author contributions All authors contributed to designing the study, analyzing the data, and writing the manuscript. All authors approved the final version of the manuscript.

Compliance with ethical standards

Conflict of interests The authors declared that they had no conflicts of interest with respect to their authorship and publication of this article.

References

- Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience*, *7*, 66–80.
- Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, *20*, 723–742.
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–624.
- Block, R. A., & Gruber, R. P. (2014). Time perception, attention, and memory: A selective review. *Acta Psychologica*, *149*, 129–133.
- Dunn, T. L., Lutes, D. J., & Risko, E. F. (2016). Metacognitive evaluation in the avoidance of demand. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 1372–1388.
- Dunn, T. L., & Risko, E. F. (2016). Toward a metacognitive account of cognitive offloading. *Cognitive Science*, *40*(5), 1080–1127.
- Eisenberger, R. (1992). Learned industriousness. *Psychological Review*, *99*, 248–267.
- Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, *47*, 381.
- Gray, W. D., & Fu, W. T. (2004). Soft constraints in interactive behavior: The case of ignoring perfect knowledge in-the-world for imperfect knowledge in-the-head. *Cognitive Science*, *28*, 359–382.
- Gray, W. D., Sims, C. R., Fu, W. T., & Schoelles, M. J. (2006). The soft constraints hypothesis: A rational analysis approach to resource allocation for interactive behavior. *Psychological Review*, *113*, 461.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs: Prentice Hall.
- Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology: General*, *139*, 665–682.
- 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*, 661–679.
- Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. New York: Wiley.
- Naparstek, S., & Henik, A. (2010). Count me in! On the automaticity of numerosity processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 1053–1059.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97–113.
- Rosenbaum, D. A. (2008). Reaching while walking: Reaching distance costs more than walking distance. *Psychonomic Bulletin & Review*, *15*, 1100–1104.
- Rosenbaum, D. A. (2017). *Knowing hands—The cognitive psychology of manual control*. New York: Cambridge University Press.
- Rosenbaum, D. A. (2012). The tiger on your tail: Choosing between temporally extended behaviors. *Psychological Science*, *23*, 855–860.
- Rosenbaum, D. A., Brach, M., & Semenov, A. (2011). Behavioral ecology meets motor behavior: Choosing between walking and reaching paths. *Journal of Motor Behavior*, *43*, 131–136.
- Rosenbaum, D. A., Carlson, R. A., & Gilmore, R. O. (2001). Acquisition of intellectual and perceptual-motor skills. *Annual Review of Psychology*, *52*, 453–470.
- Rosenbaum, D. A., Chapman, K. M., Coelho, C. J., Gong, L., & Studenka, B. E. (2013). Choosing actions. *Frontiers in Psychology*, *4*, 273, <https://doi.org/10.3389/fpsyg.2013.00273>
- 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.
- Rosenblum, L. D. (2010). *See what I'm saying: The extraordinary powers of our five senses*. New York: Norton.
- Sainburg, R. L. (2005). Handedness: Differential specializations for control of trajectory and position. *Exercise and Sport Sciences Reviews*, *33*, 206–213.
- Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*, *3*, 207–214.
- Siegler, R. S., & Lemaire, P. (1997). Older and younger adults' strategy choices in multiplication: Testing predictions of ASCM using the choice/no-choice method. *Journal of Experimental Psychology: General*, *126*, 71–92.
- Stevens, S. S. (1961). The psychophysics of sensory function. In W. A. Rosenblith (Ed.), *Sensory communications* (pp. 1–33). Cambridge: MIT Press.
- Van Der Wel, R. P., & Rosenbaum, D. A. (2007). Coordination of locomotion and prehension. *Experimental Brain Research*, *176*, 281–287.
- Westbrook, A., & Braver, T. S. (2015). Cognitive effort: A neuroeconomic approach. *Cognitive, Affective, & Behavioral Neuroscience*, *15*, 395–415.
- Westbrook, A., Kester, D., & Braver, T. S. (2013). What is the subjective cost of cognitive effort? Load, trait, and aging effects revealed by economic preference. *PLOS ONE*, *8*, e68210
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*, 625–636.
- Zakay, D., & Block, R. A. (1996). The role of attention in time estimation processes. *Advances in Psychology*, *115*, 143–164.