People Underestimate their Capability to Motivate Themselves without Performance-based Extrinsic Incentives

Kei Kuratomi[1], Laura Johnsen[2], Shinji Kitagami[3], Aya Hatano[4], Kou Murayama[5]

1. Faculty of Education, Gifu Shotoku Gakuen University, Japan

2. Department of Psychology, University of Reading, United Kingdom

3. Department of Psychology, Nagoya University, Japan

4. Department of Cognitive Psychology in Education, Graduate School of Education, Kyoto University, Japan

5. Hector Research Institute of Psychology and Education, University of Tübingen, Germany

Correspondence concerning this article should be addressed to Kei Kuratomi, Faculty of Education, Gifu Shotoku Gakuen University, 1-1, Takakuwanishi, Yanaizu-cho, Gifu-shi, Gifu, Japan. E-mail: kkuratomi@gifu.shotoku.ac.jp and Kou Murayama, Hector Research Institute of Education Sciences and Psychology, University of Tübingen, Europastraße 6, 72072 Tübingen, Germany. E-mail: k.murayama@uni-tuebingen.de

This research was supported in part by the Marie Curie Career Integration Grant, Award (CIG630680); Leverhulme Trust Research Leadership Award (RL-2016-030); JSPS KAKENHI (Grant Numbers 15H05401; 16H06406, 18H01102; 18K18696) and Jacobs Fellowship (all to K. Murayama).

Abstract

Research has shown that we are endowed with a remarkable capacity to motivate ourselves in the absence of extrinsic incentives (i.e. intrinsic motivation). However, little research has been conducted to investigate whether we accurately appreciate the power of intrinsic motivation. The current research aimed to examine the metacognitive accuracy of the extent to which people can motivate themselves without performance-based extrinsic incentives. Participants were presented with a relatively long and repetitive task without extrinsic incentives, and before doing the task, they were asked to predict their motivation on completion of the task. Across seven experiments using a variety of tasks with different populations from different countries, participants were consistently engaged in the task more actively than they predicted. When participants were provided with performance-based monetary rewards, however, this bias was diminished. These results indicate that we tend to underappreciate our capability to sustain our motivation without extrinsic incentives.

Key words: metamotivation, metacognition, affective forecasting, interest.

Word count: 148 words

Sustaining and enhancing motivation and engagement have been a long-standing issue in many applied settings such as education and work places. One of the primary sources of motivation is extrinsic rewards, which refer to the incentives outside of, and instrumental to, the activity itself, such as monetary incentives (Kruglanski, 1975). However, humans are also endowed with a remarkable capacity to sustain motivation *without* being fueled by extrinsic rewards (Deci & Ryan, 1985; Renninger & Hidi, 2016; see also Murayama, FitzGibbon, & Sakaki, 2019). Even if one comes across a situation where they need to work on a relatively repetitive task, without extrinsic rewards, the person often exhibits a variety of self-regulation or self-control skills to sustain their motivation. For example, Sansone, Weir, Harpster, and Morgan (1992) showed that when participants were asked to perform necessary, but uninteresting activities (e.g. copying a letter matrix), they self-motivated themselves by generating strategies to make the task more engaging (e.g. self-setting a challenging goal).

But do we really know how intrinsically motivated we are? The main purpose of the current study is to examine people's metacognitive accuracy of their intrinsic motivation --- the capacity to motivate themselves without performance-based extrinsic incentives. Previous research has examined various factors and strategies that influence people's self-regulation of motivation, providing some important insights into how we can promote/undermine intrinsic motivation (Duckworth, Gendler, & Gross, 2014; Hulleman & Harackiewicz, 2009; Thoman, Sansone, & Pasupathi, 2007; Wolters, 2003; Yeager et al., 2014). Little research however, has examined the metacognitive accuracy of people's metacognitive belief about the power of intrinsic motivation. This is unfortunate, as it is our metacognitive belief that guides our decisions and strategies to regulate motivation; if we have inaccurate metacognitions about our motivation, we are likely to make suboptimal decisions to regulate one's or others' motivation

(Dickson & Wendorf, 1999; Heath, 1999; MacGregor, 1960). Murayama, Kitagami, Tanaka, and Raw (2017) called such metacognitive belief or awareness about our motivation, *metamotivation* (see also Scholer & Miele, 2017; Sholer, Miele, Murayama & Fujita, 2018), and pointed out the possibility that inaccurate metamotivation can lead people to adopt ineffective motivating strategies, despite their well-meant intentions of enhancing motivation. If we underappreciate our capacity of intrinsic motivation, for example, we may end up spending more money than necessary to motivate others.

Although sparse, there have been several studies that examined people's belief about motivation (Gurland & Glowacky, 2011; Heath, 1999; Miller & Ratner, 1998; Murayama et al., 2017; Woolley & Fishbach, 2015). These studies mainly focused on people's metamotivational belief about extrinsic rewards, and they generally suggested that people tend to overestimate the motivating power of extrinsic incentives, despite potential cultural differences (DeVoe & Iyengar, 2004). None of the existent studies, however, have highlighted people's belief about their intrinsic motivation in the absence of performance-based extrinsic incentives. In addition, most of these studies simply examined people's belief about motivation only, and did not directly compare people's belief with actual motivation (for an exception, see Woolley and Fishbach, 2015). This makes it difficult to evaluate the accuracy/inaccuracy of people's metacognitive belief about how motivation works.

The current research provides empirical evidence to aid understanding of people's metamotivational belief about intrinsic motivation and its accuracy. In seven experiments, using various tasks, with various populations, participants worked on a relatively repetitive and long task, without performance-based extrinsic incentives. Critically, to examine the accuracy of participants' metamotivation, before doing the task, participants were asked to make a prediction about their motivation at the end of the task. This predicted motivation was then compared with participants' actual motivation, assessed after the task. We expected participants to underestimate the power of intrinsic motivation --- that participants would enjoy and get engaged in the task, without performance-based incentives, more actively than they predict. As indicated earlier, humans have notable capacity to motivate and engage themselves without extrinsic incentives by producing self-motivating strategies or intrinsic rewards (Sansone et al., 1992; Renninger & Hidi, 2016). However, in a recently proposed reward-learning framework of intrinsic motivation (or interest), Murayama (2019; see also Murayama et al., 2019) argued that such self-motivating mechanisms are generative and invisible in nature, and it is difficult for people to conceive of them when they make a prediction about their own motivation. As a result, participants may exhibit significant underestimation of their actual motivation.

Experiment 1

Experiment 1 aims to provide the initial evidence that people underestimate their power of intrinsic motivation.

Method

Participants. A total of 50 UK adults were recruited from Scientific Prolific (21 female, M_{age} = 24.4)[1]. In this and the following studies, no interim statistical tests were conducted. The experiments were approved by the Ethics Committees of the University of Reading (Experiment 1, 2, 4 and 5) and Nagoya University (Experiment 3).

Measures. In this and the following studies, we assessed participants' intrinsic motivation for the respective task with two scales used in Elliot and Harackiewicz (1996): task engagement and task enjoyment[2]. Both task engagement (e.g., “I concentrated on the activity”) and task enjoyment (e.g., “I enjoy doing [task name]”) were assessed with three items on a 7-point scale.

Procedure. The experiment was conducted online. Participants were instructed to work on a series of association production tasks for approximately 20 minutes. In this task, participants were presented with words in a random order and asked to produce as many words as possible that were associated with the presented word. Participants typed in these words on a screen within 20 seconds before proceeding to the next trial. There were 60 words in total.

Participants first completed 1-min practice trials and rated their intrinsic motivation for the task (task engagement and task enjoyment) so that they gained a sense of how the task felt to complete. Before the main task, participants were asked to make a prediction about their intrinsic motivation at the end of the 20-min task (“Now you will do the same task for 20 minutes. Please make a prediction about how you would feel about the task in 20 minutes”). After the main task, participants were asked to rate their current (actual) intrinsic motivation for the task.

Results. The comparison between the predicted motivation and actual motivation showed that participants enjoyed (M = 4.15, SD = 1.47) and were engaged in (M = 5.35, SD = 0.90) the task more than they predicted (M = 3.31, SD = 1.58; M = 4.77, SD = 1.14), ts (49) = 4.68 and 3.67, ps < .05, ds = 0.66, and 0.52. These results indicate that, consistent with our expectation, participants underestimated their actual motivation during the task.

Experiment 2

The purpose of Experiment 2 is to replicate the findings with a different task and with a between-subjects design where we manipulated the prediction and actual motivation as a between-subject factor. In addition, for exploratory purpose, we manipulated task difficulty to address the possibility that the findings from Experiment 1 were caused by the relatively high task performance (i.e., participants generally felt that they were doing well, which might have increased their task engagement).

Method

Participants. A total of 83 (73 female, M_{age} = 19.6) students at a UK university participated in the study. Participants were randomly assigned to one of the four conditions in a factorial 2 (Prediction: predicted motivation and actual motivation) X 2 (Task Difficulty: easy and difficult) design. The majority of participants completed the experiment in a group session with two participants.

Measures. We employed the same measures of task engagement (3 items) and task enjoyment (3 items) as Experiment 1.

Task. A total of 16 easy lists and 16 difficult lists of word-ordering task (one for the practice task and the other 15 for the main task) were used in a paper-and-pencil format. Each list was comprised of eight words. Participants were presented with a word list and had one minute to re-write the list in alphabetical order on their task sheet. Easy word lists typically included words that were simple to differentiate from each other: For example, they had different letters at the start of each word; *anterior, barracks* and *chewed*. Conversely, for the difficult word lists, the majority of words shared the same first and some subsequent letters, making them more difficult to differentiate; *hasten, hastily* and *hatched*.

Procedure. On arriving in the lab room, participants were given instructions about the task. Participants then completed a one-minute practice trial of the word-ordering task, at the difficulty level they had been assigned, followed by self-reported motivation questions about the practice trial. In the main task, participants were first informed that they would go through 15 word lists and they would have one minute for each word list. Participants in the predicted motivation condition were asked to make a prediction about their task engagement and task enjoyment after the main task, whereas participants in the actual motivation condition rated their current task engagement and task enjoyment after the main task[3].

Results

To examine whether the task difficulty manipulation was successful, a 2 (Prediction: predicted motivation and actual motivation) X 2 (Task Difficulty: easy and difficult) between-subjects ANOVA was conducted on task performance (i.e. the number of correctly sorted words). The main effect of Task Difficulty was significant, $F(1, 79) = 143.6$, $p < .01$, $\eta_G^2 = .65$, indicating that task performance in the Easy condition was higher ($M = 7.44$, $SD = 0.74$) than that in the Difficult condition ($M = 4.30$, $SD = 1.54$). Neither of the main effect of condition nor the interaction effect was statistically significant, *ps* = 0.05 and 0.86, respectively.

To test our hypothesis, the same 2 X 2 ANOVA was conducted on the task engagement and task enjoyment. Results showed, as illustrated in Figure 1, that participants' task engagement exhibited a significant main effect of Prediction, $F\ (1, 79) = 18.06$, $p < .01$, $\eta_G^2 = .19$, indicating that participants' prediction about task engagement (for Easy condition, $M = 4.08$, $SD = 1.47$; for Difficult condition, $M = 4.47$, $SD = 1.02$) was lower than the actual task engagement (for Easy condition, $M = 5.20$, $SD = 1.00$; for Difficult condition, $M = 5.53$, $SD = 1.12$). Neither of the main effect of Task Difficulty nor the interaction effect was significant, *ps* = .17 and .91, respectively. These findings replicated the results of Experiment 1. On the other hand, for task enjoyment, none of the effects were statistically significant, $\eta_G^2 = 0.00 – 0.02$, *ps* > .21.

Experiment 3a

Experiments 1 and 2 showed that people become intrinsically-motivated more than they predicted (especially in terms of task engagement) without performance-based extrinsic incentives. However, it is possible that these findings do not specifically reflect people's tendency to underestimate their power of intrinsic motivation without performance-based extrinsic incentives, but simply indicate people's general tendency to underestimate all types of motivated behaviors (see Scholer & Miele, 2016 for the importance of considering different types of motivation in examining metamotivational belief). To tease apart these two possibilities, Experiment 3 manipulated the provision of performance-based extrinsic rewards. If the previous results were caused, at least in part, by people's underestimation of intrinsic motivation (without performance-based extrinsic incentives), the underestimation is expected to be reduced when people make a prediction about their motivated behavior which is contingent upon performance-based extrinsic incentives.

Experiment 3a also sought to replicate and extend the previous findings in two respects. First, we assessed participants' motivation not only with self-reported questions, but also with a task that can objectively quantify task engagement. Second, we also manipulated the time duration of the task to explore whether and how the accuracy of metamotivation changes as a function of task length.

Method

Participants. A sample of 168 (86 female, M_{age} = 18.5) students at a Japanese university were recruited, after excluding (prior to data analysis) a participant for whom the experimenter incorrectly timed the experiment and a participant who was disengaged in the course of the experiment. Participants were randomly assigned to one of eight conditions in a factorial 2

(Reward: no reward and reward) X 4 (Time: 1min, 5min, 10min, and 20min) design. Motivation prediction was manipulated within subjects. The experiment was run individually in a lab room.

Measures. We employed the same measures of task engagement (3 items) and task enjoyment (3 items) as Experiments 1 and 2. These questions were rated on a 5-point scale in this experiment.

Task. We used a self-regulatory task reported in Baumeister, Bratslavsky, Muraven, and Tice (1998). Participants were provided with papers of academic articles written in English, and told to cross off all instances of the letter *e* within the time limit. For the practice task consisting of three trial runs (30 seconds per trial), we prepared three short paragraphs. For the main task (1min, 5min, 10min, or 20min), we prepared eight papers full of words. For both the practice trials and the main task, the number of words is so large that, when participants made a prediction, it is very unlikely that they took into account the possibility of finishing all the materials within the time limit.

The task was very simple and straightforward; meaning participants' task engagement should directly influence the performance of the task. In other words, the performance of this task (i.e. the number of letters *e* crossed off within the time limit) can be considered to be linearly related to the amount of participants' active engagement in the task. Accordingly, we used the number of letter *e*'s crossed off as an alternative index of task engagement.

Procedure. Participants were instructed about the task, and worked on three 30-second practice trials, followed by self-reported questions about the trials. For these practice trials, participants were provided with the feedback on how many letters participants correctly crossed off for each 30-second trial. Participants were then told about the time limit of the main task, depending on their experimental conditions (1min, 5min, 10min or 20min). Participants in the reward condition were further told that they would obtain 1 Japanese yen (approximately 1 cent) for each letter *e* that they correctly crossed off.

Before the main task, participants were asked to make a prediction about their task engagement by (1) rating their predicted task engagement and task enjoyment at the end of the experiment and (2) indicating how many letters they thought they would cross off within the given time limit. To facilitate the accurate calibration of their basic performance for the task, participants were allowed to look back at the feedback they obtained for the practice trials. After making the prediction, participants worked on the main task, and upon finishing the task, they rated their current task engagement and task enjoyment.

Results

In the following analysis, one participant who made an unrealistically high prediction about task performance (more than 5SD above the mean across participants of the same task duration) was eliminated. To examine the effects of reward and task duration on the accuracy of metamotivation, we conducted a linear mixed-effects model (Murayama, Sakaki, Yan, & Smith, 2014) with Prediction (effect coded; -1 = actual motivation; 1 = predicted motivation), Reward (effect coded; -1 = no reward; 1 = reward), Time (continuous variable; 1, 5, 10, and 20), and their two-way and three-way interactions as the predictors of task engagement, task enjoyment, and task performance. Participants were treated as a random effect. The analysis was conducted with Mplus (Muthen & Muthen, 2004).

For task engagement, consistent with the previous experiments, the results revealed a significant main effect of Prediction, $B = -0.11$, $SE = 0.03$, $p < .01$, indicating that participants generally underestimated their self-reported task engagement for the main task. Importantly, this effect was qualified by the critical Prediction X Reward interaction, which was marginally significant, $B = 0.054$, $SE = 0.03$, $p = .077$. Simple main effect analysis showed that participants underestimated task engagement in the no reward condition, $B = -0.165$, $SE = 0.04$, $p < .01$, whereas this effect was not significant in the reward condition, $B = -0.06$, $SE = 0.04$, $p = .27$. To visualize the nature of the observed Prediction X Reward interaction effect and the other effects, Figure 2 plotted the predicted values of task engagement for each condition. As is clear from the Figure 2, overall, the underestimation of task engagement is observed in the no reward condition, but it is less clear in the reward condition. This pattern was especially clear in the 20 min condition. The slight differences in the pattern of the Prediction X Reward as a function of Time may be produced by the combination of the other main effects and the interactions (although some of them were not statistically significant).

For task enjoyment, consistent with task engagement, the results showed a significant main effect of Prediction, $B = -0.12$, $SE = 0.06$, $p < .05$, indicating that participants generally underestimated their self-reported task enjoyment; hence they enjoyed it more than they predicted they would. However, this effect was not qualified by higher-order interactions (*ps* > .21), indicating that the manipulation of reward did not have a significant influence on the underestimation of task enjoyment.

The same analysis was applied to task performance (i.e. objective measure of task engagement). Again, the results exhibited a significant main effect of Prediction, $B = -21.1$, $SE = 6.1$, $p < .01$, indicating that participants generally underestimated their task performance. Importantly, this effect was qualified by a significant Prediction X Reward X Time three-way interaction, $B = 1.7$, $SE = 0.9$, $p < .05$. Simple interaction analysis showed that, in the no reward condition, participants generally underestimated their task performance, $B = -30.2$, $SE = 9.2$, $p < .01$ and this underestimation was magnified as the duration of the task increased, $B = -3.4$, *SE*

$= 1.3$, $p < .01$. However, in the reward condition, neither the overall tendency of underestimation, nor the change in the pattern as a function of Time was observed, *ps* > .15. To visualize the nature of this three-way interaction, Figure 3 plots the predicted values of task performance. Consistent with the simple interaction analysis, the underestimation of motivation was generally observed in the no reward condition, especially when task duration is longer, but this pattern was not observed in the reward condition.

Experiment 3b

The aim of Experiment 3b is to replicate the findings that people underestimate their task engagement even with an objective measurement of motivation (i.e. letters crossed off) in the no reward condition. Furthermore, we extended and examined the robustness of the prior study by additionally investigating the effects of prior warnings about the underestimation of motivation. Several studies have suggested that people's metacognitive inaccuracy is rather persistent and resistant even after they receive verbal warnings (Yan et al., 2016). We were interested in whether metamotivational inaccuracy also shows this resistance to verbal warnings.

Participants. A sample of sixty-four (forty-six female, age information not collected) university students at a Japanese university were recruited. Participants were randomly assigned to one of the two conditions: a warning condition or a control condition. Motivation prediction was manipulated within subjects (i.e. same participants were measured both on predicted and actual motivation with self-reported measures and objective index of task engagement). The experiment was conducted in a lab room.

Measures. We employed the same self-reported measures of task engagement (3 items) and task enjoyment (3 items) as Experiment 3a (Cronbanch's αs = .72-.82). In addition to these measures, the experiment included 16 items which were prepared as filler items.

Tasks. The task was identical with that used in Experiment 3b, expect that the current study only used the 5-minutes version in the no reward condition.

Procedure. Participants were instructed about the task and worked on three 30-second practice trials. For these practices, participants were provided with feedback each time so that they can accurately calibrate their task performance (i.e. number of letter *e* crossed off) when making a prediction.

In the warning condition, participants were provided with the following warning before making the prediction: "Several recent studies, including Deci et al. (2015) and Murayama et al. (2017), have shown that participants tend to mispredict their motivation and performance for the task. Specifically, the studies have found that we tend to rate the predictive motivation and performance lower than the actual motivation and performance after the task. Simply put, many people think that they would not be that motivated to work on the task before doing it, but in fact, they were more motivated than their prediction. Based on these findings, please try to predict as accurately as possible how you will feel after the task is completed."

Before the main task, participants were asked to predict their motivation after the main task in terms of ratings of intrinsic motivation/task engagement and task performance (i.e. number of letter *e* crossed off). Participants then performed the main task. After the completion of this task, participants again rated their current task engagement and task enjoyment.

Results. Prior to conducting main data analysis, we excluded two participants due to procedural errors, and two other participants whose task performance was more than 2SD from the mean. As a result, there were 30 participants in each of the conditions.

A 2 (Prediction: predicted motivation and actual motivation) X 2 (Waring: warning and control) ANOVA was conducted on the self-reported task engagement and task enjoyment. Replicating the previous findings, the ANOVA for task engagement showed a significant main effect of Prediction, F (1, 58) = 13.15, p < .001, η_G^2 = .05, indicating that participants underestimated the task engagement (M = 3.80, SD = 0.81) in comparison to the actual task engagement (M = 4.12, SD = 0.66). Although the task engagement was numerically higher in the waring condition (M = 4.12, SD = 0.78) than in control condition (M = 3.80, SD = 0.69), the main effect of warning did not reach significance, F (1, 58) = 3.73, p = .060, η_G^2 = .05. Also, the interaction effect was not significant, F (1, 58) = 0.14, p = .71, η_G^2 = .00. Analysis of task enjoyment also showed a significant main effect of Prediction, F (1, 58) = 15.27, p < .001, η_G^2 = .04. This result indicated that participants underestimated the task enjoyment (M = 3.12, SD = 0.76) in comparison to the actual task enjoyment (M = 3.43, SD = 0.86). The other main or interaction effects were not statistically significant, η_G^2 = .00-.01, ps > .33.

To examine the difference between the predicted and actual task performance (i.e. behavioral assessment of motivation), a 2 (Prediction: predicted task performance and actual task performance) X 2 (Waring: warning and control) ANOVA was conducted. The main effect of Prediction was significant, F (1, 58) = 58.87, p < .001, η_G^2 = .30, indicating that participants underestimated the task performance (M = 161.1, SD = 71.65) in comparison to the actual task performance (M = 231.1, SD = 27.40). The other main or interaction effects were not statistically significant, η_G^2 = .01-.03, ps > .09.

Experiments 4a and 4b

In Experiments 4a and 4b, we aimed to replicate the findings of Experiment 3a with different tasks.

Participants. A total of 95 UK adults were recruited from Scientific Prolific for Experiment 4a. An additional 101 UK adults were recruited from the same platform in Experiment 4b. However, we excluded participants whose error rate was more than 20% prior to the main data analysis, resulting in 80 participants for Experiment 4a (27 female, M_{age} = 24.2) and 80 participants for Experiment 4b (35 females, M_{age} = 24.0). For each experiment, participants were randomly assigned to one of the 2 (Prediction: predicted motivation and actual motivation) X 2 (Reward: no reward and reward) conditions. The experiments were run online and were part of a larger study examining the relationship between cognitive demand and effort avoidance (Kuratomi, Shigemasu, & Murayama, in prep). Motivation assessments were included in these experiments to specifically test the hypothesis of the current research.

Measures. We employed the same measures of task engagement (3 items) and task enjoyment (3 items) as Experiment 3 (Cronbach's αs = .59-.92).

Tasks. In Experiment 4a, participants worked on a flanker task and in Experiment 4b, participants worked on a number judgement task. In the flanker task, participants were presented with one the four letter-strings for 100 ms --- NNNNN, ZZZZZ, ZZNZZ, NNZNN. Their task was to identify the central letter of the string (ignoring the other letters) as "N" or "Z" by pressing a button. In the number judgement task, participants were presented with a number between 1 and 9 (excluding 5) and their task was to make a judgement depending on the color of the number. Specially, when the number was in red, participants were required to judge whether the number was odd or even by pressing a key (i.e., parity judgement). When the number was in blue, participants were required to judge whether the number was smaller or larger than 5 (i.e., magnitude judgement). Key and color assignments were counterbalanced across participants. In both experiments, before seeing a stimulus, participants were presented with two different card decks and asked to select one of the decks. Unbeknownst to the participants, one card deck was associated with higher task demands (i.e., participants were likely to see a mixed-letter string or more frequent switch of the color or the number) than the other card deck. This experimental manipulation was made to address the main purpose of the larger research project describe earlier, but irrelevant to the current research. For both tasks, there was one block of 400 trials and the task length was about 30 minutes.

Procedure. Participants were first given instructions about the main task (flanker task in Experiment 4a or number judgement task in Experiment 4b), and completed the practice trials (20 trials for Experiment 4a and 30-60 trials for Experiment 4b depending on the task performance of the practice trials) and self-reported motivation questions about the practice trials. In the main task, participants were informed of the length of the task (30 min), and participants in the reward condition were further told that they would earn £ 0.01 for every 2 successful trials. Participants in the predicted motivation condition were then asked to make a prediction about their task engagement and task enjoyment after the main task. On the other hand, participants in the actual motivation condition rated their current task engagement and task enjoyment only after finishing the main task.

Results

Error rate for the task was generally low (flanker task = 7.3%; task-switching task = 10.4%), indicating that these participants were generally focused on the task. A 2 (Prediction: predicted motivation and actual motivation) X 2 (Reward: no reward and reward) X 2 (Task: flanker and task-switch) between-subjects ANOVA on task performance (error rate) revealed the main effects of task, indicating that error rate for the task-switch task was larger than that for the flanker task, F (1, 152) = 21.00, p < .001, η_G^2 = .12. Neither of the main effect of condition nor the interaction effect was statistically significant, ps > .28.

To test our hypothesis, the same 2 (Prediction: predicted motivation and actual motivation) X 2 (Reward: no reward and reward) conditions x 2 (Task: flanker task and number judgement task) ANOVA was conducted on the task engagement and task enjoyment. Replicating the previous findings, results for task engagement showed a significant main effect of Prediction, F (1, 152) = 21.82, p < .01, η_G^2 = .13, indicating that participants underestimated the task engagement (M = 2.30, SD = 0.88) in comparison to the actual task engagement (M = 2.86, SD = 0.65). The main effect of the reward was also significant, F (1, 152) = 4.38, p < .05, η_G^2 = .03. Critically, these main effects were qualified by a significant Prediction X Reward interaction, F (1, 152) = 5.29, p < .05, η_G^2 = .03. As can be seen in Figure 4, whereas participants substantially underestimated task engagement in the no reward condition, F (1, 152) = 24.3, p < .01, η_G^2 = .14, that underestimation was reduced in the reward condition, and it was no longer significant, F (1, 152) = 2.81, p = .10, η_G^2 = .02. These results indicate that participants underestimate their motivation especially in the absence of performance-based extrinsic incentives. The other main or interaction effects were not statistically significant, η_G^2 = 0.00 – 0.03, ps > .23.

Results for task enjoyment also showed a significant main effect of Prediction, F (1, 152) = 23.45, p < .01, η_G^2 = .13. These results mirror the results of the task engagement, suggesting that participants underestimated the task enjoyment (M = 2.05, SD = 0.92) in comparison to the actual task enjoyment (M = 2.73, SD = 0.99). The main effect of Reward was also statistically significant, F (1, 152) = 21.21, p < .01, η_G^2 = .12. However, unlike task engagement (and consistent with Experiment 3), the critical Prediction X Reward interaction was not statistically significant, F (1, 152) = 0.00, p = .98, η_G^2 = .00. Reward X Task interaction was significant, suggesting that the positive effect of Reward was larger in the flanker task (M = 2.02, SD = 0.77 for the no-reward condition; M = 3.01, SD = 1.14 for the reward condition) than in the number judgement task (M = 2.13, SD = 0.86 for the no-reward condition; M = 2.43, SD = 0.98 for the reward condition). The other main or interaction effects were not statistically significant, η_G^2 = .00-.02, ps > .09.

Replication of Reward Effect (Experiment 5)

In Experiments 3a, 4a, and 4b, we showed that people underestimated how engaging the task would be without performance-based incentives; but this effect was weaker when participants were provided with performance-based extrinsic incentives. However, in Experiment 3a, the critical Prediction X Reward interaction was only marginally significant (p = .077; although a similar effect was observed in task performance). In Experiment 4a and 4b, the Prediction X Reward interaction was significant, but the data from two experiments were combined. To demonstrate the robustness of our finding, we conducted an exact replication of Experiment 4a[4]. Experiment 4a was selected because it showed the largest effect size for the Prediction X Reward interaction effect.

Participants. In Experiment 4a, the effect size of the critical interaction effect between prediction and reward was 0.32 in Cohen's f metric. Using this effect size estimate, 79 participants were required to achieve the statistical power of 80%. With this number in mind, we collected as much data as possible within budgetary limits. As a result, 127 participants were recruited for this replication study from Scientific Prolific, and after applying the same exclusion criteria as with Experiment 4a (determined *a priori*), 106 participants were included in the main data analysis (64 female, M_{age} = 25.3). The experiment was conducted online and participants were randomly assigned to one of the 2 (Prediction: predicted motivation and actual motivation) X 2 (Reward: no reward and reward) conditions.

Measures. We employed the same measures of task engagement (3 items) and task enjoyment (3 items) as Experiment 3 (Cronbach's αs = .43 - .88).

Tasks and Procedure. The experiment employed the same flanker task used in Experiment 4a. Procedure was also identical with Experiment 4a.

Results

Error rate for the task was generally low (7.6%), indicating that these participants were generally focused on the task. A 2 (Prediction: predicted motivation and actual motivation) X 2 (Reward: no reward and reward) between-subjects ANOVA on task performance (error rate) revealed no significant main effects or interaction effect (*ps* > .46).

To test our hypothesis, 2 (Prediction: predicted motivation and actual motivation) X 2 (Reward: no reward and reward) ANOVA was conducted on the task engagement and task enjoyment. Replicating our main finding, task engagement demonstrated a significant main effect of Prediction, F (1, 102) = 10.20, $p < .01$, $\eta_G^2 = .09$, suggesting that participants again underestimated the task engagement when making a prediction (M = 3.50, SD = 0.74) in comparison to the actual task engagement (M = 3.90, SD = 0.62). Importantly, this main effect was qualified by a significant Prediction X Reward interaction, F (1, 102) = 5.21, $p < .05$, $\eta_G^2 = .05$. As indicated in Figure 5, participants underestimated task engagement in the no reward condition, F (1, 102) = 14.16, $p < .01$, $\eta_G^2 = .12$, but this underestimation was substantially diminished in the reward condition, F (1, 102) = 0.44, $p = .51$, $\eta_G^2 = .00$. The main effect of reward was not statistically significant, F (1, 102) = 0.01, $p = .93$, $\eta_G^2 = .00$.

Unlike the previous experiment, task enjoyment also showed a similar pattern. Specifically, 2 X 2 ANOVA showed a significant main effect of Prediction, F (1, 102) = 15.90, p < .01, η_G^2 = .13, suggesting the overall underestimation of task enjoyment (M = 2.34, SD = 0.86 for the predicted task enjoyment; M = 2.99, SD = 0.97 for the actual task enjoyment). The main effect of reward was also significant, F (1, 102) = 6.13, p < .05, η_G^2 = .0.06. Critically, Prediction X Reward interaction also exhibited a significant effect, F (1, 102) = 4.80, p < .05, η_G^2 = .04. Simple main effect analyses revealed that participants in the no reward condition significantly underestimated task enjoyment, F (1, 102) = 18.01, p < .01, η_G^2 = .15 (M = 1.90; SD = 0.66 for the predicted task enjoyment; M = 2.96, SD = 0.92 for the actual task enjoyment) but this was not observed in the reward condition, F (1, 102) = 1.72, p = .19, η_G^2 = .02 (M = 2.70; SD = 0.83 for the predicted task enjoyment; M = 3.01, SD = 1.04 for the actual task enjoyment).

General Discussion

Across seven experiments using various tasks with different populations from different countries, we consistently observed that participants underestimated the extent to which they would be intrinsically motivated towards a task without performance-based extrinsic incentives. In other words, participants were engaged in the task more actively without performance-based incentives than they predicted. These results indicate the inaccuracy of our metacognition about our ability to regulate and sustain intrinsic motivation.

In the literature of metacognition, researchers have repeatedly found that we often make an inaccurate prediction about their future learning (Kornell & Bjork, 2008; Murayama, Blake, Kerr, & Castel, 2016; Rhodes & Castel, 2008; Soderstrom & Bjork, 2015). The literature of affective forecasting also shows that people routinely mispredict their future emotion and well-being (Wilson & Gilbert, 2005; see also Loewenstein, Weber, Hsee, & Welch, 2001; Schwarz, 2015). Our research expanded those findings, showing that such metacognitive inaccuracy is also observed in people’s predictions or beliefs about motivation (i.e. metamotivation). The problem of inaccurate metacognition is that people are likely to engage in suboptimal decision making, because people’s decision making is largely based on their metacognitive belief about what is right. In fact, previous studies showed that inaccurate metacognition inadvertently promotes the use of maladaptive learning strategies when people study (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013). Other studies also revealed that misprediction of future affect makes people fail to maximize their happiness (e.g., Gilbert & Ebert, 2002). As the current research found that people underestimate the power of intrinsic motivation, the logical next step for future research is to examine the impact of such inaccurate metamotivation on people’s actual self- or other-motivating behaviors. From the current findings, we can reasonably expect that people rely on extrinsic incentives more than necessary to motivate themselves or others.

Another next important step is to examine the mechanisms underlying this underestimation effect. There are several possibilities. For example, participants underestimated the actual intrinsic motivation because they were unable to appreciate their capability to generate strategies to make the task engaging (Sansone et al., 1992). It is also possible that people did not appreciate the power of automation. Specifically, participants were not able to understand that the physical and mental cost that participants perceived in the practice would diminish over time; because participants become more efficient at completing the task, the task becomes less demanding. Another likely mechanism is that people overlooked the potential rewarding feelings that they acquired by becoming proficient in the task. Critically, these explanations are not mutually exclusive and it is possible that they jointly result in the underestimation effect (Murayama, 2019). Future studies should systematically manipulate task properties to examine how and when people underestimate task engagement without performance-based extrinsic incentives.

It is worth noting that we consistently found the underestimation of intrinsic motivation for task engagement, but the results were less consistent for task enjoyment. This observation suggests that our findings do not simply reflect the inaccuracy of the affective forecasting (Wilson & Gilbert, 2005) and cannot be fully explained by the factors underlying affective forecasting (e.g., emotional adaptation; Ubel, Loewenstein, & Jepson, 2003). Research on interest argues that task engagement and task enjoyment are essential components of interest, but represent different phases of interest development (Hidi & Renninger, 2006; Sansone & Thoman, 2005). Specifically, the literature suggests that initial (often extrinsic) task engagement precedes the development of task enjoyment, which requires more internalization of task values (see also Renninger & Hidi, 2002). As participants worked on the tasks for only up to 30 minutes in the current experiments, and participants were not provided with sufficient reasons to work on the task, it is likely that they did not develop interest to the extent they could internalize and enjoy the task. Although we did not specifically predict the different pattern of results between task engagement and task enjoyment, the current findings provided some evidence for the importance of distinguishing these elements of task interest (see also Goldsmith & Dhar 2013).

As indicated in the introduction, several previous studies examined people's metamotivational belief on extrinsic incentives, and suggested that people tend to overestimate the motivational power of extrinsic incentives (Heath, 1999; Miller & Ratner, 1998; Woolley & Fishbach, 2015). Although the present study was not specifically designed to examine the accuracy of the motivational power of extrinsic incentives, a close scrutiny of the results exhibits a pattern consistent with these interpretations. That is, in Experiments 3a, 4, and the replication experiment, the predicted motivation substantially increased from the no-reward condition to the reward condition but the actual increase in motivation was much smaller or rewards even slightly decreased motivation (see Figures 2-5). This pattern was supported by the statistically significant Prediction X Reward interactions reported in the results (see Woolley & Fishbach, 2015 for similar pattern of results). These observations indicate that people overestimate the positive effects of extrinsic rewards on motivation. This finding poses an interesting contrast with the underestimation of the power of intrinsic motivation.

Over the past decades, a vast amount of research has been conducted to examine the nature, determinants, and consequences of intrinsic motivation, cultivating and advancing our understanding of human motivated behavior (Ryan & Deci, 2017). Substantial body of research in cognitive psychology has also revealed that our metacognitive belief about cognitive process (e.g., memory) is often inaccurate, endangering our self-regulated behavior (Bjork, Dunlosky, & Kornell, 2013). However, despite the prominence of both lines of work, and despite its critical role in applied settings such as education, organization, and policy making, surprisingly little research has integrated these perspectives, examining the metacognitive accuracy of intrinsic motivation. We hope that the current findings open a new avenue for a number of future research programs, elucidating more detailed characteristics of our belief in how intrinsically-motivated we are.

References

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), 1252-1265. http://dx.doi.org/10.1037/0022-3514.74.5.1252

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology, 64*, 417-444.

Boekaerts, M. (1995). Self-regulated learning: Bridging the gap between metacognitive and metamotivation theories. *Educational Psychologist, 30*(4), 195–200. https://doi.org/10.1207/s15326985ep3004_4

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

DeVoe, S. E., & Iyengar, S. S. (2004). Manager's theories of subordinates: A cross-cultural examination of manager perceptions of motivation and appraisal of performance. *Organizational Behavior and Human Decision Processes. 93*(1), 47-61. https://doi.org/10.1016/j.obhdp.2003.09.001

Goldsmith, K., & Dhar, R. (2013). Negativity bias and task motivation: Testing the effectiveness of positively versus negatively framed incentives. *Journal of Experimental Psychology: Applied, 19*(4), 358–366. https://doi.org/10.1037/a0034415

Dickson, M. W., & Wendorf, C. A. (1999). *Managers' implicit theories of motivation: How managers think motivation "Works".* Paper presented at the 107th annual meeting of the American Psychological Association. Boston, MA.

Duckworth, A. L., Gendler, T. S., & Gross, J. J. (2014). Self-Control in School-Age Children. *Educational Psychologist, 49*(3), 199-217. http://dx.doi.org/10.1080/00461520.2014.926225

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14*(1), 4-58. http://dx.doi.org/10.1177/1529100612453266

Elliot, A. J., & Harackiewicz, J. M. (1996). Approach and avoidance achievement goals and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology, 70*, 461-475. http://dx.doi.org/10.1037/0022-3514.70.3.461

Gilbert, D. T., & Ebert, J. E. J. (2002). Decisions and revisions: The affective forecasting of changeable outcomes. *Journal of Personality and Social Psychology, 82,* 503-514. http://dx.doi.org/10.1037/0022-3514.82.4.503

Gurland, S. T., & Glowacky, V. C. (2011). Children's theories of motivation. *Journal of Experimental Child Psychology, 110*(1), 1-19. https://doi.org/10.1016/j.jecp.2011.03.013

Heath, C. (1999). On the social psychology of agency relationships: Lay theories of motivation overemphasize extrinsic incentives. *Organizational Behavior and Human Decision Processes, 78*(1), 25-62. https://doi.org/10.1006/obhd.1999.2826

Hidi, S., & Renninger, K. A. (2006). The Four-Phase Model of Interest Development. *Educational Psychologist, 41*(2), 111-127. http://dx.doi.org/10.1207/s15326985ep4102_4

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science, 326*(5958), 1410-1412. http://dx.doi.org/10.1126/science.1177067

Kornell, N., & Bjork, R. A. (2008). Learning Concepts and Categories: Is Spacing the "Enemy of Induction"? *Psychological Science, 19*(6), 585-592. http://dx.doi.org/10.1111/j.1467-9280.2008.02127.x

Kruglanski, A. W. (1975). The endogenous-exogenous partition in attribution theory. *Psychological Review, 82*(6), 387-406. http://dx.doi.org/10.1037/0033-295X.82.6.387

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin, 127*(2), 267-286. http://dx.doi.org/10.1037/0033-2909.127.2.267

Murayama, K., FitzGibbon, L., & Sakaki, M. (under review). Process account of curiosity and interest: A reward learning model of knowledge acquisition.

MacGregor, D. (1960). *The human side of enterprise* (Vol. 21): New York.

Miele, D. B., & Scholer, A. A. (2016). Self-regulation of motivation. In K. R. Wentzel & D. B. Miele (Eds.), *Handbook of Motivation at School* (2nd Edition, pp. 363–384). New York: Routledge.

Miller, D. T., & Ratner, R. K. (1998). The disparity between the actual and assumed power of self-interest. *Journal of Personality and Social Psychology, 74*(1), 53-62. http://dx.doi.org/10.1037/0022-3514.74.1.53

Murayama, K. (2019). A reward-learning framework of autonomous knowledge acquisition: An integrated account of curiosity, interest, and intrinsic-extrinsic rewards. Preprint available online.

Murayama, K., Blake, A. B., Kerr, T., & Castel, A. D. (2016). When Enough Is Not Enough: Information Overload and Metacognitive Decisions to Stop Studying Information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *94*, 914-924. http://dx.doi.org/10.1037/xlm0000213

Murayama, K., Fitzgibbon, L., & Sakaki, M. (2019). Process account of curiosity and interest: A reward learning framework of knowledge acquisition. . https://doi.org/10.31219/osf.io/hbcz5.

Murayama, K., Kitagami, S., Tanaka, A., & Raw, J. A. L. (2017). People's Naiveté About How Extrinsic Rewards Influence Intrinsic Motivation. *Motivation Science*, 2(3), 138-142. http://dx.doi.org/10.1037/mot0000040

Murayama, K., Matsumoto, M., Izuma, K., & Matsumoto, K. (2010). Neural basis of the undermining effect of monetary reward on intrinsic motivation. *Proceedings of the National Academy of Sciences*, *107*(49), 20911–20916. https://doi.org/10.1073/pnas.1013305107

Murayama, K., Sakaki, M., Yan, V. X., & Smith, G. M. (2014). Type 1 error inflation in the traditional by-participant analysis to metamemory accuracy: A generalized mixed-effects model perspective. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *40*, 1287-1306. http://dx.doi.org/10.1037/a0036914

Muthen, L. K., & Muthen, B. O. (2004). Mplus use's guide. Los Angeles, CA: Author.

Renninger, K. A., & Hidi, S. (2002). Student interest and achievement: Developmental issues raised by a case study. In A. Wigfield & J. Eccles (Eds.), *Development of achievement motivation: A volume in the educational psychology series.* (pp. 173-195). San Diego, CA: Academic.

Renninger, K. A., & Hidi, S. (2016). *The power of interest for motivation and engagement.* New York, NY: Routledge.

Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *Journal of Experimental Psychology: General, 137*(4), 615-625. http://dx.doi.org/10.1037/a0013684

Ryan, R. M., & Deci, E. L. (2017). *Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness.* Guilford Publications.

Sansone, C., & Thoman, D. B. (2005). Interest as the Missing Motivator in Self-Regulation. *European Psychologist, 10*(3), 175-186. https://doi.org/10.1027/1016-9040.10.3.175

Sansone, C., Weir, C., Harpster, L., & Morgan, C. (1992). Once a boring task always a boring task? Interest as a self-regulatory mechanism. *Journal of Personality and Social Psychology, 63*(3), 379-390. http://dx.doi.org/10.1037/0022-3514.63.3.379

Scholer, A. A., & Miele, D. B. (2016). The role of metamotivation in creating task-motivation fit. *Motivation Science, 2*(3), 171–197. https://doi.org/10.1037/mot0000043

Scholer, A. A., Miele, D. B., Murayama, K., & Fujita, K. (in press). New Directions in Self-Regulation: The Role of Metamotivational Beliefs. *Current Directions in Psychological Science*

Schwarz, N. (2015). Metacognition. In M. Mikulincer, P. R. Shaver, E. Borgida, & J. A. Bargh (Eds.), *APA handbook of personality and social psychology, Volume 1: Attitudes and social cognition* (pp. 203-229). Washington, DC, US: American Psychological Association.

Soderstrom, N. C., & Bjork, R. A. (2015). Learning Versus Performance. *Perspectives on Psychological Science, 10*(2), 176–199. https://doi.org/10.1177/1745691615569000

Thoman, D. B., Sansone, C., & Pasupathi, M. (2007). Talking about interest: Exploring the role of social interaction for regulating motivation and the interest experience. *Journal of Happiness Studies, 8*(3), 335-370. https://doi.org/10.1007/s10902-006-9016-3

Ubel, P. A., Loewenstein, G., & Jepson, C. (2005). Disability and sunshine: Can hedonic predictions be improved by drawing attention to focusing illusions or emotional adaptation? *Journal of Experimental Psychology-Applied, 11*(2), 111–123. https://doi.org/Doi 10.1037/1076-898x.11.2.111

Wilson, T. D., & Gilbert, D. T. (2005). Affective Forecasting: Knowing What to Want. *Current Directions in Psychological Science, 14*(3), 131-134. https://doi.org/10.1111/j.0963-7214.2005.00355.x

Wolters, C. A. (2003). Regulation of motivation: Evaluating an underemphasized aspect of self-regulated learning. *Educational Psychologist, 38*(4), 189-205. http://dx.doi.org/10.1207/s15326985ep3804_1

Woolley, K., & Fishbach, A. (2015). The experience matters more than you think: People value intrinsic incentives more inside than outside an activity. *Journal of Personality and Social Psychology, 109*(6), 968-982. http://dx.doi.org/10.1037/pspa0000035

Yeager, D. S., Henderson, M. D., Paunesku, D., Walton, G. M., D'Mello, S., Spitzer, B. J., & Duckworth, A. L. (2014). Boring but important: A self-transcendent purpose for learning fosters academic self-regulation. *Journal of Personality and Social Psychology, 107*(4), 559-580. http://dx.doi.org/doi:10.1037/a0037637

Footnotes

1. In all experiments, participants received some form of compensation (e.g., money, course credit) for their participation. However, these compensations were made for their participation only; compensations were not contingent upon the motivation or the quality of performance for the task.

2. Data from all experiments can be obtained in Open Science Framework (https://osf.io/pjq8m/). We also included other measures for exploratory purpose in Experiments 2 and 3 but task engagement and task enjoyment are the only measures that were commonly used across all experiments.

3. In Experiments 2 and 4a/4b (and the Replication Experiment), participants in the Prediction condition actually did the main task following the prediction, and rated their motivation after finishing the main task. This is because we needed to ensure that participants spent a certain amount of time for the experiment to justify the advertised compensation. The results of these measures (i.e. the analysis of predicted motivation and actual motivation focusing only on the participants in the predicted motivation condition) were consistent with the main findings reported in this paper.

4. This is the only exact replication study that we conducted.

Figure Captions

Figure 1

Task engagement as a function of Prediction and Task difficulty in Experiment 2. Error bars represent standard errors.

Figure 2

Self-reported task engagement as a function of Prediction, Reward, and Time in Experiment 3 (predicted values from the mixed effects model). Error bars represent standard errors, which were computed using delta method.

Figure 3

Task performance (objective measure of task engagement) as a function of Prediction, Reward, and Time in Experiment 3 (predicted values from the mixed effects model). Error bars represent standard errors, which were computed using delta method.

Figure 4

Task engagement as a function of Prediction and Reward in Experiment 4. Error bars represent standard errors.

Figure 5

Task engagement as a function of Prediction and Reward in the replication experiment of Experiment 4a. Error bars represent standard errors.

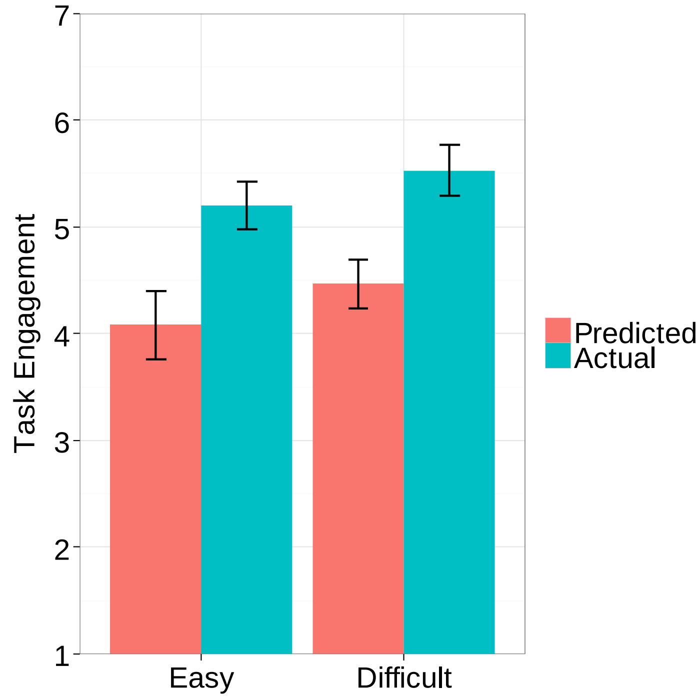

Figure 1

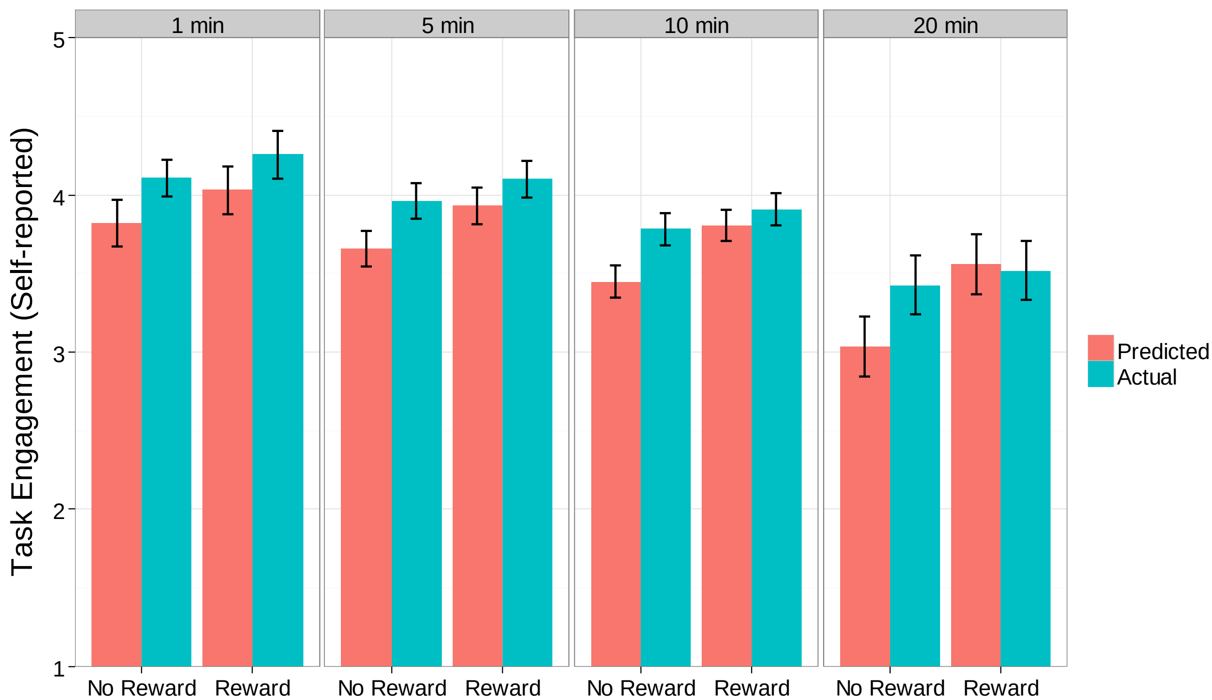
1 min
5 min
10 min
20 min
Task Engagement (Self-reported)
5
4
3
2
1
No Reward
Reward
Predicted
Actual

Figure 2

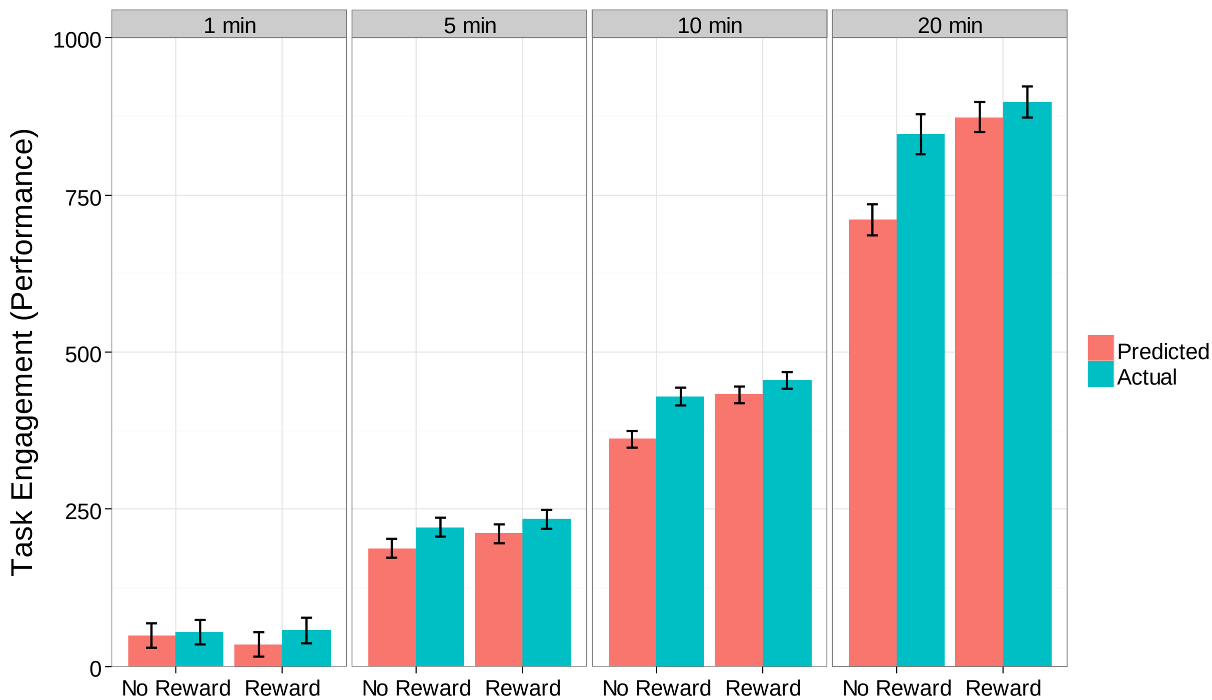
1 min
5 min
10 min
20 min
1000
750
500
250
0
Task Engagement (Performance)
No Reward
Reward
No Reward
Reward
No Reward
Reward
No Reward
Reward
Predicted
Actual

Figure 3

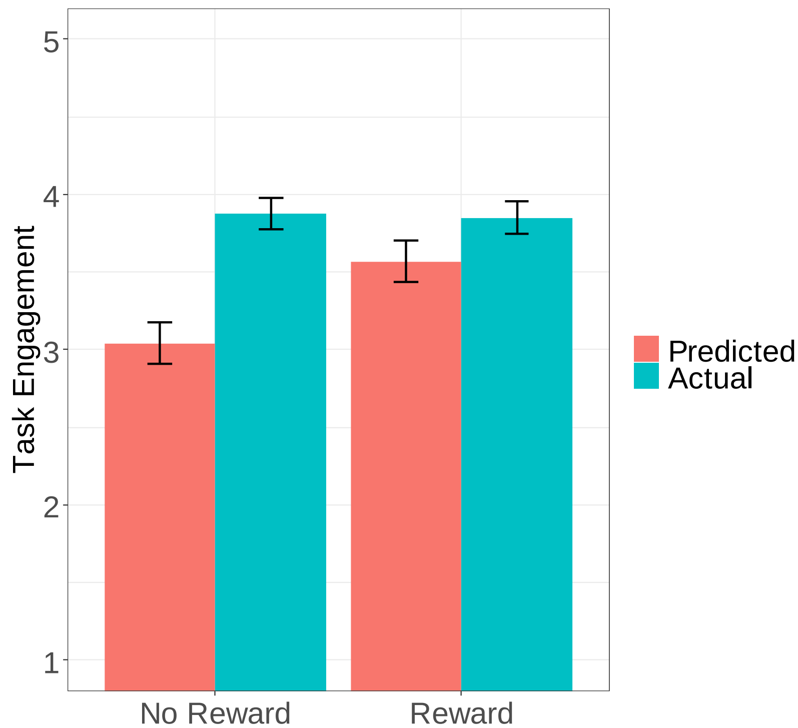

Figure 4

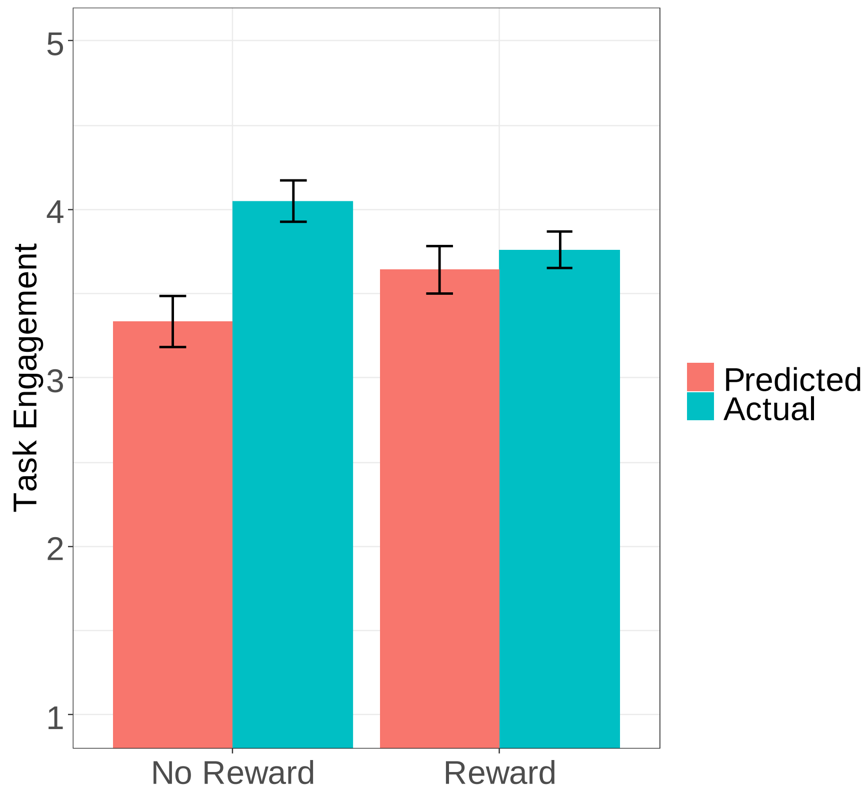

Figure 5