Does It Pay to Play? Undermining Effects of Monetary Reward and Gamification in a Web-Based Task

Alexandra N. Bitter¹, Trent Wondra², Sean M. McCrea¹, Ali Darzi³, and Vesna D. Novak^{4, 5}

¹ Department of Psychology, University of Wyoming

² Department of Psychology, Barry University

³ Department of Psychology, University of Pittsburgh

⁴ Department of Electrical Engineering, University of Wyoming

⁵ Department of Electrical Engineering and Computer Science, University of Cincinnati

The increasing use of self-paced computer and web-based platforms for e-learning and work has led to renewed interest in promoting intrinsic motivation. A classic finding in motivation literature is the undermining effect: Task interest is reduced when previously presented monetary rewards are eliminated. Efforts to prevent undermining effects on web-based tasks have focused on adding game-like features, such as points or leaderboards (i.e., “gamification”), to maintain task interest. On the other hand, some have raised the concern that these game elements could be a reward that undermines intrinsic motivation, such that task engagement is reduced when these game features are eliminated. We, therefore, conducted two preregistered studies of the undermining effects of pay and gamification in a web-based memory task. Results revealed a small undermining effect for task engagement and performance when previously presented incentives were removed from the task. Exploratory analyses suggested slightly different mechanisms behind each type of undermining effect. Finally, we observed large benefits for engagement and performance of performance-based monetary rewards, relative to adding game elements alone. Implications for explanations of the undermining effect and application to online work and learning contexts are discussed.

Keywords: intrinsic motivation, undermining effect, gamification

Supplemental materials: <https://doi.org/10.1037/tmb0000056>.sup

Intrinsic motivation is understood as the extent to which engaging in a task serves as its own reward (Deci, 1971; Kruglanski et al., 2018). A classic finding in this literature concerns the benefits and costs to intrinsic motivation when extrinsic rewards (e.g., pay or

praise) are presented for engaging in the task. Specifically, expected extrinsic rewards have been found to undermine subsequent intrinsic motivation on the task (Deci, 1971; Lepper et al., 1973). A typical paradigm involves asking participants to engage in an

Action Editor: C. Shawn Green was the action editor for this article.

ORCID iDs: Alexandra N. Bitter <https://orcid.org/0000-0002-6105-4403>; Trent Wondra <https://orcid.org/0000-0002-8052-6234>; Sean M. McCrea <https://orcid.org/0000-0003-4241-5530>; Vesna D. Novak <https://orcid.org/0000-0001-9143-2682>.

Funding: This research was supported by the National Science Foundation under Grant No. 1717705 (Domen Novak, PI, Sean M. McCrea Co-PI).

Disclosures: The authors declare that they have no conflict of interest.

Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Potential participants were informed about the study’s aims and assured that participation would be voluntary and anonymous. Informed consent was obtained electronically and responses were identified only by a code. As participating in our study did not cause harm or discomfort that went beyond everyday experiences, this procedure was in line with the guidelines of the national Society of Psychology and our university’s ethics committee.

Prior Data Use: There have been no prior uses of this data.

Data Availability: Data, analytic methods, and study materials are publicly available to other researchers. <https://data.mendeley.com/datasets/jnzyy6rc5g/draft?a=6f3a3a2d-09f2-4af3-b615-c8f9a3aff218>

Open Science Disclosures:

The data are available at <https://data.mendeley.com/datasets/jnzyy6rc5g/1>

The experiment materials are available at <https://data.mendeley.com/datasets/jnzyy6rc5g/1>

The preregistered design and analysis plan is accessible at https://aspredicted.org/BQF_LXO for Study 1 and https://aspredicted.org/PPW_JZQ for Study 2.

Open Access License: This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC-BY-NC-ND). This license permits copying and redistributing the work in any medium or format for noncommercial use provided the original authors and source are credited and a link to the license is included in attribution. No derivative works are permitted under this license.

Contact Information: Correspondence concerning this article should be addressed to Alexandra N. Bitter, Department of Psychology, University of Wyoming, 1000 E. University Avenue, Laramie, WY 82071, United States. Email: abitter@uwyo.edu

intrinsically interesting task (e.g., puzzles, pinball, or a labyrinth marble game), with or without extrinsic rewards. In a subsequent free-choice period, participants have the option of whether to engage in the task in the absence of any extrinsic rewards (Deci, 1971; Harackiewicz et al., 1984; Smith & Pittman, 1978). In these studies, the removal of previously available extrinsic rewards reduced engagement in the task during the subsequent free-choice period, relative to conditions in which the extrinsic rewards were never offered. These findings suggest that intrinsic motivation was undermined by the presentation of extrinsic reward.

A recent movement has attempted to leverage computer-based technologies to increase engagement in tasks in both work and educational contexts, without offering potentially undermining extrinsic rewards. Specifically, game-like features including points and leaderboards can be added to a task (so-called “gamification”) with the aim to make it more enjoyable and intrinsically motivating (Ferrara, 2013). However, there have been few studies to determine whether classic undermining effects of tangible rewards occur in such gamified tasks, particularly in a learning or work environment. In addition, gamification in the context of work or educational tasks may itself be viewed as a form of reward that can undermine intrinsic motivation. That is, if a task is introduced in a gamified format, interest in engaging in the same type of task may be reduced when game features are no longer available. We, therefore, sought to conceptually replicate classic undermining effects of removing performance-based pay on an online gamified task, and examine whether gamification itself undermines intrinsic motivation.

Explanations of the Undermining Effects of Reward

The most frequently invoked explanation for the undermining effect derives from Self-Determination Theory (SDT; Deci & Ryan, 1985; Ryan & Deci, 2000). According to this view, rewards will only increase intrinsic motivation to the extent that they facilitate basic needs for psychological well-being: relatedness, competence, and autonomy. Extrinsic rewards tend to signal the demands of others and are, therefore, experienced as controlling (Deci, 1971; Deci & Ryan, 1985). As a result, rewarded tasks fail to meet basic needs of autonomy and do not facilitate psychological well-being. On the other hand, to the extent that rewards serve as positive feedback concerning the development of task competence, they could serve to satisfy a core need (Arkes, 1979; Deci & Ryan, 1985; Rosenfield et al., 1980). Thus, the subjective experience of reward determines whether the provision of rewards will undermine or promote intrinsic motivation.

Consistent with this view, Hagger and Chatzisarantis (2011) found undermining effects only among those with a stronger autonomy orientation (i.e., perceived actions as originating from the self rather than by external events). A related finding is that highly salient, task-contingent rewards attenuate the predictive effects of need satisfaction on performance (Cerasoli et al., 2016). That is, need satisfaction matters less to performance when individuals understand the link between behavior and reward. This perspective is also bolstered by neuroscientific evidence. A study by Murayama et al. (2010) reported reduced subjective value resulting from the removal of performance-based rewards, indicated by decreased activation in the striatum and midbrain. These authors argued that such findings are consistent with a reduced sense of autonomy.

Other theoretical approaches suggest that extrinsic motivation can “crowd-out” intrinsic motivation. A recent example, the Structural

Model of Intrinsic Motivation (Kruglanski et al., 2018), suggests that undermining effects occur when a task is instrumental to the completion of multiple goals (Kruglanski et al., 2018; Shah & Kruglanski, 2000). Added goals reduce the relative importance of (i.e., crowd out) the other goals served by engaging in the task (Zhang et al., 2007). The upshot of such a process is that overall task interest will be reduced if the task can no longer fulfill an associated goal (Kruglanski et al., 2018). For example, if a task is associated with the goals to earn money and to have fun, subsequent interest will be reduced if one can no longer earn money (or no longer have fun) by engaging in the task. Whereas engaging in the task originally fulfilled two goals, now, it only fulfills one. In contrast to the predictions of SDT, this perspective holds that removing any kind of instrumental reward will reduce subsequent task interest.

Reviews of the Undermining Effect

There has been considerable debate about the robustness and boundary conditions of the undermining effect. At least nine meta-analyses of the literature have been conducted, often arriving at different conclusions due to investigator choices in how to classify studies by methodology (Cameron et al., 2001; Cameron & Pierce, 1994; Deci et al., 1999, 2001; Eisenberger et al., 1999; Eisenberger & Cameron, 1996; Rummel & Feinberg, 1988; Tang & Hall, 1995; Wiersma, 1992). For example, Deci et al. (1999) report relatively pervasive negative effects of expected tangible reward, whereas Cameron et al. (2001) report finding undermining effects on free-choice behavior only when these rewards were contingent on doing well overall or on the number of items solved. Both analyses suggested that effects are more consistently negative on free-choice behavioral measures than on self-report measures, and for tasks that were initially intrinsically interesting. Cameron et al. (2001) argue that rewards can actually have beneficial effects for intrinsic motivation and interest when expressed as verbal praise and for tasks that are initially less interesting (see also Lepper, 1998).

An additional set of issues concern how applicable these findings are to work and learning environments. Some have argued that studies of undermining effects should involve rewards tied to achievement (Cameron et al., 2005) and presented over the course of the task (Hidi, 2016; Reiss, 2005). In a study following these procedures, Cameron et al. (2005) found evidence that performance-contingent rewards actually increased student’s intrinsic motivation for a puzzle-solving task.

Finally, it should be noted that many of the studies considered by prior reviews were underpowered by current standards (Button et al., 2013; Curran-Everett, 2017). This could have affected estimates of the overall size of the undermining effect and the importance of identified moderators. Additional work in this area could, therefore, strengthen conclusions regarding the conditions that lead to an undermining effect.

Gamification

A recent approach to increasing task engagement while avoiding the undermining effect involves the “gamification” of learning tasks (also, serious games, game-based learning; Laine & Lindberg, 2020). The basic notion is that computer-based e-learning tasks or work-related tasks can be made more intrinsically interesting by adding game-like elements, such as levels, points, and cooperative or competitive play. Tasks that are more intrinsically motivating should not require

extrinsic rewards for engagement, thereby reducing deleterious undermining effects over the long term (Deterding et al., 2011; Wouters et al., 2013). For example, a repetitive arm-motor task intended for physical rehabilitation can be made more intrinsically motivating by having participants control a paddle in a game of pong against a computer or human opponent (Novak et al., 2014), or labeling photos can be made more interesting by awarding points and displaying a leaderboard (Mekler et al., 2017). Past work has generally focused on tasks in educational contexts, finding that participants perform better when learning with a gamified task (e.g., Domínguez et al., 2013).

Game elements or mechanics intended to increase intrinsic interest are rich and varied. These include setting specific goals, providing immediate and visible feedback, use of adaptive difficulty, social aspects of competition or cooperation, freedom of choice for the player, multiple trials or rounds to gain proficiency, and the use of storytelling and fantasy (Dicheva et al., 2015). Particularly for more long-term learning contexts (e.g., learning a language or trigonometry), game features can be quite complex or even use existing games to convey educational concepts (Laine & Lindberg, 2020). A thorough consideration of the impact of these different features and learning contexts is beyond the scope of the current work. We focus here on the effects of a few very basic game features, notably the provision of points for completing task-relevant actions. Awarding points is among the most popular mechanisms of educational games that have been studied (Dicheva et al., 2015). We also focus on a relatively simple task, comparable to those typically studied in the undermining literature (see Cameron et al., 2001; Deci et al., 1999) and similar to those used in popular online “brain-training” programs (e.g., BrainHQ, Luminosity).

Despite findings that it can improve task performance, actual evidence that gamification increases intrinsic motivation is weaker (Wouters et al., 2013). Few studies have utilized valid measures of intrinsic interest or statistically compared gamified to nongamified systems (paid or unpaid) on these measures (Seaborn & Fels, 2015). Some have also raised the concern that popular game elements involving points, badges, and leaderboards may be viewed as a type of controlling or extrinsic reward that would undermine intrinsic interest in the task (Amriani et al., 2013; Hanus & Fox, 2015). For example, Amriani et al. (2013) found that removing gamification (specifically, points, badges, a leaderboard, and visualized goal progress) undermined participation in an e-learning task. Prior reviews of undermining effects (e.g., Cameron et al., 2001; Deci et al., 1999) differentiate between tangible and verbal rewards, and it is unclear whether the provision of game elements such as points function more like the former or the latter.

More research is, therefore, needed to examine whether gamification can be experienced as controlling or intrinsically motivating. If game features such as the provision of points are experienced as controlling, they may lead to undermining effects (Deci et al., 1999). Removing gamification could also undermine subsequent task interest due to reduced task instrumentality, as described by the Structural Model of Intrinsic Motivation (Kruglanski et al., 2018).

Present Studies

The present work had several goals. First, we sought to conceptually replicate classic research on undermining effects (Deci, 1971; Lepper et al., 1973) with a gamified web-based computer task. Specifically, we examined the effects of expected, tangible, performance-contingent rewards on free-choice behavior on a

web-based training task that was initially enjoyable. This approach advances past work by examining the undermining effect in a more applied context in which the participant is given an explicit goal of learning or skill improvement on the task (see also Cameron et al., 2005).

Second, we sought to test whether undermining effects are observed when removing game features from this type of task. If gamification in this context is viewed as controlling, SDT (Deci, 1971; Ryan & Deci, 2000) would predict that these game features undermine intrinsic motivation. It is also possible that removing gamification reduces the instrumentality of the task for fulfilling the goal of having fun, decreasing subsequent task interest (Kruglanski et al., 2018; Zhang et al., 2007). We focused on several basic game elements that have previously been found (e.g., Amriani et al., 2013) to undermine intrinsic interest, namely, the provision of feedback and the awarding of points. These game features are directly comparable to the provision of performance-contingent monetary rewards that supposedly lead to undermining effects (Deci, 1971).

Third, we assessed task enjoyment, perceived competence, tension, effort, extrinsic motivation, and instrumentality of the task for making money and having fun. This allowed us to determine whether any undermining effects we observed were due to rewards being viewed as controlling or reducing instrumentality of the task for other goals (Kruglanski et al., 2018; Ryan & Deci, 2005).

Finally, the design of these studies allowed for an overall comparison of the effectiveness of gamification and performance-based pay for task engagement and performance. That is, independent of the effects of removing a reward, one might expect greater task engagement with performance-based pay than with gamification among online workers.

We designed a location memory task that varied in whether gamification, performance-based pay, or both types of rewards were offered. The study was realized in four conditions (see Table 1). To examine the undermining effects of pay, we compared a condition in which performance-based pay was presented in the initial phase and then removed in the free-choice phase to a condition in which performance-based pay was never offered (see also Deci, 1971). In the pay never offered condition, participants engaged in a gamified version of the task in both the initial phase and the free-choice phase. This condition provides a baseline for task engagement when pay was never present for participants. In the pay removal condition, participants first engaged in a version of the task that had both gamification and performance-based pay. In the free-choice phase, the task was still gamified, but the performance-based pay was removed.

A parallel set of conditions provided a test of the undermining effects of gamification. We compared a condition in which gamification was presented in the initial phase and then removed in the free-choice phase to a condition in which gamification was never offered. In the gamification never offered condition, participants engaged in a version of the task with performance-based monetary rewards in both the initial phase and the free-choice phase. This condition provided a baseline for task engagement when gamification was never present for participants. In the gamification removal condition, participants first engaged in a version of the task that had both gamification and performance-based monetary reward. In the free-choice phase, the monetary reward was still available, but the game features were removed.

Table 1
Task Features in the Initial and Free-Choice Phase by Experimental Condition

Type of effect	Reward offered	Initial phase	Free-choice phase
Undermining pay	Pay never offered Pay removed	Gamified, no pay Gamified plus pay	Gamified, no pay Gamified, no pay
Undermining gamification	Gamification never offered Gamification removed	Nongamified, pay Gamified plus pay	Nongamified, pay Nongamified, pay

Within each type of undermining condition, the reward offered manipulation provides a conceptual replication of the classic undermining effect on task engagement and performance.

This design also allows us to determine the relative levels of engagement and performance to be expected from offering performance-based pay or gamification for online workers. Specifically, in the undermining pay conditions, gamification is always present in both phases, whereas in the undermining gamification conditions, performance-based pay is always present in both phases.

Study 1

We recruited a large enough sample to detect small-to-moderate undermining effects, consistent with the size of undermining effects observed in the literature (Cameron et al., 2001; Deci et al., 1999). We also preregistered the study design, hypotheses, and analysis plan (see Preregistered Material). Materials (i.e., image files and JSPsych code), data, and analysis script are made available.

Hypotheses

Our primary hypotheses concerned behavior in the free-choice phase. We based our hypothesis on SDT and past work demonstrating the undermining effect of pay (Deci, 1971; Ryan & Deci, 2000). According to SDT, undermining effects should only occur when incentives are experienced as controlling. We, therefore, expected reduced task engagement and performance during the free-choice behavior period when pay had been removed, compared to when pay was never offered. This would reflect a significant effect of reward offered within the undermining pay condition. Because we designed the gamified task to be enjoyable, we did not expect gamification to have an undermining effect. Specifically, engagement and performance were not expected to differ when gamification was removed compared to when gamification had never been offered. Thus, we did not expect an effect of reward offered within the undermining gamification condition. Combined, these predictions result in an expected Type of effect $\times$ Reward offered interaction on the number of item blocks completed and score in the free-choice period.

With regard to the manipulation checks taken after the completion of the initial phase, we predicted the following. Effort and instrumentality for improving memory were expected to be high across all conditions. Critical tests of the manipulation checks involve comparisons within the reward never offered conditions, as the initial phase tasks are identical in the removal conditions (see Table 1). We expected enjoyment, competence, and instrumentality for having fun to be maximized in the pay never offered condition, relative to the gamification never offered condition. Conversely, we expected

pay during the initial phase to increase, and gamification during the initial phase to decrease, extrinsic motivation. Thus, we expected extrinsic motivation, tension, and instrumentality for making money to be maximized in the gamification never offered condition, relative to the pay never offered conditions.

Finally, we expected performance in the initial phase would be improved by performance-contingent pay relative to gamification alone. Thus, we expected a significant interaction, such that performance during the initial phase would be worse in the pay never offered condition than in the gamification never offered condition, with no difference in the removal conditions. Pay and gamification are both present during the initial phase of the removal conditions and should not differ.

Method

Participants and Design

Participants were recruited from Amazon's Mechanical Turk (MTurk) in return for \$0.40 payment and the possibility of a bonus for performance. Participants had to be located in the U.S. and have a 95% acceptance rate on a minimum of 100 previous Human Intelligence Tasks (HITs). They were also told prior to starting the study that they would need a number-pad on their keyboard. Participant recruitment was conducted with CloudResearch Turk-Prime (Litman et al., 2017), which additionally blocked anyone who previously completed the study or shared an IP address with a previous participant.

The sample size was determined in advance using G*Power (Faul et al., 2007) for a classic undermining effect of moderate size (Cohen's $d = .50$, independent groups t test, $1 - \beta = .80$, and $\alpha = .05$). These parameters yielded a recommended sample size of $N = 128$, which was then quadrupled to $N = 512$ to account for the full design and expected null undermining effect of gamification (Giner-Sorolla, 2018). By quadrupling the sample size, we have 80% power to detect a classic undermining effect of pay of $d = .36$ or larger (independent groups t test, $\alpha = .05$). For comparison, this is roughly commensurate with the estimated average size of undermining effects in studies with expected, tangible rewards overall ($d = .36$), or specifically performance-contingent rewards ($d = .28$), as reported by Deci et al. (1999).

Consistent with this plan, we collected data until we had $N = 512$ who entered a completion code, with a minimum of $n = 115$ per cell. We recorded all instances in which the study was started. Of the $N = 938$ who opened the study, $N = 571$ completed the full study (i.e., began the free-choice behavior phase). The proportion of incomplete responses did not differ by condition, $\chi^2(3) = 5.41, p = .144$.

Participants were randomly assigned to one condition in a 2 (Type of effect: undermining pay vs. undermining gamification) $\times$ 2

(Reward offered: never offered vs. removed) between-participants design. Data analysis did not proceed until data collection was completed.

Location Memory Task

The format of the location memory task was varied to involve gamification, performance-contingent pay, or both (see also Table 1). The task was analogous to a spatial N -back task (Green et al., 2005), but was noncontinuous (i.e., memory was only assessed on certain trials) to simplify the bonus-pay structure. The task progressed from relatively easy (1-back trials) to difficult (3-back trials), though participants were not informed of this ordering. The task and all measures were implemented in JSPsych v.6 (de Leeuw, 2015). Examples of all versions of the task are provided here: <http://food.gear.host/undermining.htm>. Early versions of the task were piloted with small groups of participants to ensure that they understood the task and that the gamified version was enjoyable.

In all versions of the task, participants had to track the location of a target in a 3×3 grid (see Figure 1). A grid of neutral stimuli was presented for 1,000 ms, followed by a trial in which a target appeared in one location of the grid for 1,000 ms, and then the neutral grid for another 500 ms. Locations of the grid were numbered to correspond to the number pad of a keyboard (i.e., bottom row, from left to right: 1, 2, 3; middle row from left to right: 4, 5, 6; etc.). Location of the target on a given trial was determined on a random-with-replacement basis. The sequence of locations presented in a particular set of trials varied randomly between 3 and 6 trials in length. At the end of the sequence, participants were presented with the neutral grid and prompted to indicate the location of the target n -trials ago (one, two, or three). They had 3,000 ms to

respond before the computer advanced to the next sequence. Participants completed three sequences with a 1-back prompt, followed by three sequences with a 2-back prompt, followed by three sequences with a 3-back prompt. The task, therefore, was composed of nine total location memory judgments. In the initial phase, participants completed all nine sequences. In the free-choice phase, participants could quit the task at any point. A key aspect of performance on this type of working memory task is to attend to and encode the target locations for later retrieval (Green et al., 2005; Unsworth & Spillers, 2010). Gamification, therefore, centered on adding game elements to facilitate encoding and motivate retrieval using (a) novel, cartoon-like targets, (b) a clear goal to indicate the location of each target by pressing the corresponding key, (c) immediate visual feedback for correctly indicating the location of the target, (d) provision of points for correctly indicating the location of the target, and (e) provision of additional points for correct memory responses. We chose not to use badges or leaderboards, as these game elements could evoke social comparison processes.

Nongamified, Pay Version

In this version of the task, participants had to track the location of an orange circle that appeared in a grid of blue circles (see Figure 1). A performance contingent-reward was provided for correct memory responses: 1 cent for each correct 1-back decision, 2 cents for each correct 2-back decision, and 3 cents for each correct 3-back decision. A running tally of bonus cents earned was presented at the bottom of the screen, and total bonus earned was presented at the end of the task.

Gamified, No Pay Version

In this version of the task, participants had to track the location of a gopher in a 3×3 grid of dirt mounds (see Figure 2). Whenever a

Figure 1
Example Sequence for Nongamified, Pay Task

Figure 2*Example Sequence for the Gamified, No-Pay Task and Gamified Plus Pay Task*

gopher appeared at a given location, participants could press the corresponding key to “hit” the gopher. A successful response was indicated by the presentation of an explosion at the location for 500 ms, and 10 points were awarded. An incorrect response or no response was indicated by the presentation of the grid of mounds for 500 ms and no award of points.

In addition, participants earned points for correct location memory responses. They earned 10 points for correct 1-back location judgments, 20 points for correct 2-back location judgments, and 30 points for correct 3-back location judgments. A running tally of total points was provided during the task, and total score presented at the end of the task.

Gamified Plus Pay Version

This version of the task was similar in appearance to the gamified, no pay version. However, participants earned points and bonus pay for correct memory responses. They earned 10 points and a 1-cent bonus for each correct 1-back decision, 20 points and a 2-cent bonus for each correct 2-back decision, and 30 points and a 3-cent bonus for each correct 3-back decision. A running tally of points earned and bonus earned was provided during the task, with totals presented at the end of the task.

Manipulation Checks

The effectiveness of the manipulations in the initial phase was assessed with the following self-report measures.

Self-Reported Intrinsic and Extrinsic Motivation

Participants completed a 22-item Intrinsic Motivation Inventory (Ryan & Deci, 2005). Subscales of the measure assessed interest and enjoyment (e.g., “While playing the memory game, I was thinking about how much I enjoyed it,” Cronbach’s $\alpha = .87$), perceived competence (e.g., “I think I am pretty good at the memory game,” Cronbach’s $\alpha = .88$), tension (e.g., “I felt very tense during the memory game,” Cronbach’s $\alpha = .77$), and effort intensity (e.g., “I put a lot of effort into the memory game,” Cronbach’s $\alpha = .83$). Intrinsic motivation and need fulfillment should be reflected by higher interest and enjoyment, higher levels of competence, and lower levels of tension. Additional four items were added to the measure to directly assess extrinsic motivation (e.g., “I am only interested in playing this memory game for a reward,” Cronbach’s $\alpha = .70$). There were two attention check items (e.g., please select 3) embedded in this questionnaire. All responses were made on a seven-point scale (1 = *not at all true* to 7 = *very true*).

Instrumentality

Participants completed a three-item measure of instrumentality of the task for different goals. The items assessed instrumentality of the task for improving memory, having fun, and earning money. Each goal and the memory task were presented as a series of seven Venn diagrams ranging from no overlap to close overlap (Aron et al., 1992; Kruglanski et al., 2018). Higher scores, therefore, indicate a greater instrumentality for achieving the specified goal (i.e., greater fusion between the goal and the task).

Demographics

Participants provided their age and gender in open-ended responses.

Procedure

Participants provided informed consent prior to starting the study. They were randomly assigned to one of the four conditions (see Table 1). All participants were told that the purpose of the study was to examine tasks that could be used to train memory performance and that they had the goal of improving their memory. Participants were additionally told that they either had the goal to earn money (gamification never offered condition), have fun (pay never offered condition), or to earn money and have fun (pay removal and gamification removal conditions).

Participants assigned to the gamification never offered condition read the task instructions and then completed a brief test of understanding. Participants were asked to indicate whether the following statements were true or false: “I will not earn a bonus for remembering where the orange dot was 1, 2, or 3 trials ago”; “If I remember where the orange dot was 2 trials ago, I earn 2 cents bonus”; “If the orange dot was at locations 5, 7, 2, 3: when I am asked where it was three trials ago, I should respond with 7.” Correct answers were subsequently provided to ensure that the participants comprehended the instructions (respectively: false, true, true). The computer then presented the orange circle in each possible location in a random order, and participants were prompted to click a button when they were ready to start the task. Participants then completed nine-blocks of the nongamified, pay task.

Participants assigned to the pay never offered condition read task instructions and then completed a brief test of understanding. Participants were asked to indicate whether the following statements were true or false: “I will not earn points for remembering where the gopher was 1, 2, or 3 trials ago”; “If I remember where the gopher was 2 trials ago, I earn 20 points”; “If the gopher was at locations 5, 7, 2, 3: when I am asked where it was three trials ago, I should respond with 7.” Correct answers were subsequently provided to ensure that the participants comprehended the instructions (respectively, false, true, true). The computer then presented the gopher in each possible location in a random order, and participants were prompted to click a button when they were ready to start the task. Participants then completed nine-blocks of the gamified, no pay task.

Participants assigned to the pay removal and gamification removal conditions read task instructions and then completed a brief test of understanding. Participants were asked to indicate whether the following statements were true or false: “I will not earn points or a bonus for remembering where the gopher was 1, 2, or 3 trials ago”; “If I remember where the gopher was 2 trials ago, I earn 20 points and a 2 cent bonus”; “If the gopher was at locations 5, 7, 2, 3: when I am asked where it was three trials ago, I should respond with 7.” Correct answers were subsequently provided to ensure the participants comprehended the instructions (respectively, false, true, true). The computer then presented the gopher in each possible location in a random order, and participants were prompted to click a button when they were ready to start the task. Participants in both removal conditions then completed nine-blocks of the gamified plus pay task.

Following the initial phase, participants completed the manipulation check measures and provided demographic information.

Next, participants were told that they had the opportunity to do the memory task again for as long as they wanted, pressing the escape key at any time to quit. Instructions concerning the version of the task for the free-choice phase were provided. Those in the pay never offered and gamification never offered conditions were told the task would be the same as before. Those in the pay removal condition were told that there was no longer any monetary bonus for remembering the locations of the gophers. Those in the gamification removal condition were told that they were now to remember the location of orange circles instead of gophers and there would be no points to be earned by “hitting” the targets.

Those assigned to the pay never offered and pay removal conditions then completed the gamified, and no pay version of the task until they pressed the escape key or they completed all nine sequences. Those assigned to the gamification never offered and gamification removal conditions completed the nongamified, pay version of the task until they pressed the escape key, or they completed all nine sequences.

Participants provided a completion code to receive payment.

Results

Data Preparation

All analyses were conducted with R (R Core Team, 2020) using the Rcmdr package (Fox, 2005). We preregistered several criteria for data exclusion. Consistent with this plan, we excluded data from 11 participants who failed both attention checks, and from 54 participants who failed to respond to any block of the location memory tasks. No data needed to be excluded on the basis of providing the same response to all items of the intrinsic motivation inventory. Preregistered exclusions did not differ by condition, $\chi^2(3) = 6.21, p = .10$.

We also conducted analyses to address possible concerns of data quality. Although performance on the test of understanding the instructions was generally high, a number of participants did poorly either due to confusion or lack of attention. We, therefore, conducted the primary analyses when excluding participants who scored less than 50% on the test of understanding the instructions. This resulted in exclusions of additional 65 participants, leaving $N = 441$ (Age: $M = 39.35, SD = 12.32$; Gender: Male = 218, Female = 219, Nonbinary = 2, Did Not Identify = 2). These exclusions did not differ by condition, $\chi^2(3) = 4.27, p = .24$. There were no condition differences on understanding of the instructions ($p = .23$). A sensitivity analysis with this sample size suggested 80% power to detect small-to-moderate undermining effects within each reward offered condition ($ds > .38$). As these additional exclusions were not included in the preregistered analysis plan, we report the results under the planned analyses in Supplemental Materials A–D. Correlations, grand means, and standard deviation of all study variables are presented in Table 2. Controlling for age or gender of participants did not alter the results, and so, we do not discuss these variables further.

Manipulation Checks

Manipulation check measures assessed at the end of the initial task phase were submitted to 2 (Type of effect: undermining pay vs. undermining gamification) $\times$ 2 (Reward offered: never

Table 2
Correlations for Study Variables

Variable	1	2	3	4	5	6	7	8	9	10
1. Memory	—									
2. Fun	.532**	—								
3. Money	.306**	.199**	—							
4. Enjoy	.325**	.563**	.081	—						
5. Effort	.116*	.094*	.108*	.396**	—					
6. Competence	.175**	.311**	.207**	.385**	-.087	—				
7. Tension	-.027	-.171**	-.111*	-.163**	-.016	-.473**	—			
8. Extrinsic	-.079	-.169**	.312**	-.262**	-.039	.065	.111*	—		
9. Attempts	.145**	.097*	.017	.026	.086	-.020	.092	-.056	—	
10. Performance	.107*	-.025	.029	-.028	.145**	-.051	.086	-.012	.833**	—
<i>M</i>	4.77	4.19	4.91	4.99	5.84	3.90	3.69	4.59	4.12	4.87
<i>SD</i>	1.91	2.06	1.98	1.38	1.08	1.53	1.35	1.35	3.72	5.96

* $p < .05$. ** $p < .01$.

offered vs. removed) between-participants ANOVAs (see Table 3). Because the most critical test of the manipulations involves comparisons within the reward never offered conditions, simple effects of type were conducted within reward offered condition (see Table 4).

No significant differences were found for effort and instrumentality of improving memory, though the interaction was marginal. Means on both measures were high, indicating that individuals were generally engaged by the task and believed that the task was valid.

Analyses revealed a significant Type of effect $\times$ Reward offered interaction on instrumentality for having fun. Those in the pay never offered (gamified) condition reported higher instrumentality for having fun than did individuals in gamification never offered (pay) condition. A significant Type of effect $\times$ Reward offered interaction was also found for enjoyment. Enjoyment was lowest in the gamification never offered condition. Means were above the mid-point in both the pay removal and pay never offered conditions. Thus, conditions with gamification in the initial phase resulted in adequate levels of enjoyment and fun as measured by self-report.

The predicted interaction on extrinsic motivation was not significant. Means on this measure were slightly above the mid-point, which could suggest that participants were responding on the basis of being paid to complete the study rather than on the

performance-based bonus. Likewise, there was no interaction effect on instrumentality for making money. We did find a marginally significant Type of effect $\times$ Reward offered interaction in the expected direction on tension. Tension tended to be higher among individuals in the gamification never offered condition than in the pay never offered condition. No significant interaction effect was found for competence, though means were also in the expected direction. Finally, there were no significant effects on initial phase performance ($F_s < 1$, $p_s > .325$).

Hypothesis Tests on Free Choice Behavior

Number of blocks attempted during the free-choice phase was submitted to a 2 (Type of effect: undermining pay vs. undermining gamification) $\times$ 2 (Reward offered: never offered vs. removed) between-participants ANOVA (see Table 3). There were large effects of undermining type, such that participants attempted far more items when in the undermining gamification conditions (in which performance-contingent pay was provided) than in the undermining pay conditions (in which gamification was provided). Providing performance-contingent pay thus had large benefits for free-choice engagement relative to gamification. There was no overall effect of reward offered. We next followed the preregistered analysis plan to

Table 3
ANOVA Results for Manipulation Checks, Trial Attempts, and Performance

Variable	Reward offered				Type of effect				Interaction			
	<i>F</i> (1, 437)	<i>p</i>	<i>d</i>	95% CI _d	<i>F</i> (1, 437)	<i>p</i>	<i>d</i>	95% CI _d	<i>F</i> (1, 437)	<i>p</i>	<i>d</i>	95% CI _d
Memory	0.78	.378	0.08	[-0.10, 0.27]	0.31	.579	0.05	[-0.13, 0.24]	3.02	.083	0.17	[-0.02, 0.35]
Fun	4.92	.027	0.21	[0.02, 0.40]	4.74	.030	0.21	[0.02, 0.40]	16.35	<.001	0.39	[0.20, 0.58]
Money	3.81	.052	0.19	[0.00, 0.37]	0.78	.378	0.08	[-0.10, 0.27]	0.42	.516	0.06	[-0.12, 0.25]
Enjoy	4.76	.030	0.21	[0.02, 0.40]	4.90	.027	0.21	[0.02, 0.40]	5.38	.021	0.22	[0.03, 0.41]
Effort	0.21	.643	0.04	[-0.14, 0.23]	0.67	.412	0.08	[-0.11, 0.27]	1.11	.292	0.10	[-0.09, 0.29]
Competence	0.27	.601	0.05	[-0.14, 0.24]	0.91	.341	0.09	[-0.10, 0.28]	2.29	.131	0.14	[-0.04, 0.33]
Tension	0.47	.493	0.07	[-0.12, 0.25]	1.04	.309	0.10	[-0.09, 0.29]	3.43	.065	0.18	[-0.01, 0.36]
Extrinsic	0.43	.512	0.06	[-0.12, 0.25]	0.24	.625	0.05	[-0.14, 0.23]	0.37	.545	0.06	[-0.13, 0.25]
Attempts	0.99	.321	0.10	[-0.09, 0.28]	103.87	<.001	0.98	[0.78, 1.17]	2.82	.094	0.16	[-0.03, 0.35]
Performance	2.44	.119	0.15	[-0.04, 0.34]	109.74	<.001	1.00	[0.80, 1.20]	1.72	.190	0.13	[-0.06, 0.31]

Table 4*Independent Samples t-Tests of Type of Effect, by Reward Condition*

Variable	Pay never offered		Gamification never offered		df	t	p	d	95% CI _d
	n = 108		n = 107						
	M	SD	M	SD					
Memory	4.96	1.83	4.75	2.04	210.14 ^a	0.81	.417	0.11	[−0.16, 0.38]
Fun	4.56	1.99	3.37	1.98	213	4.40	<.001	0.60	[0.33, 0.87]
Money	4.70	2.20	4.75	2.02	213	−0.15	.879	−0.02	[−0.29, 0.25]
Enjoy	5.14	1.37	4.55	1.50	213	3.01	.003	0.41	[0.14, 0.68]
Effort	5.96	1.03	5.77	1.09	213	1.33	.184	0.18	[−0.09, 0.45]
Competence	4.04	1.60	3.68	1.51	213	1.69	.092	0.21	[−0.04, 0.50]
Tension	3.55	1.45	3.92	1.31	213	−1.96	.051	−0.27	[−0.54, 0.00]
Extrinsic	4.63	1.35	4.64	1.37	213	−0.08	.935	−0.01	[−0.28, 0.26]

	Pay removed		Gamification removed		df	t	p	d	95% CI _d
	n = 116		n = 110						
	M	SD	M	SD					
Memory	4.48	1.94	4.90	1.82	224	−1.67	.097	−0.22	[−0.48, 0.04]
Fun	4.22	1.99	4.57	2.08	224	−1.32	.188	−0.18	[−0.44, 0.09]
Money	4.95	1.88	5.24	1.77	224	−1.19	.237	−0.16	[−0.42, 0.10]
Enjoy	5.12	1.31	5.13	1.28	224	−0.08	.937	−0.01	[−0.27, 0.25]
Effort	5.81	1.09	5.83	1.10	224	−0.16	.870	−0.02	[−0.28, 0.24]
Competence	3.89	1.45	3.97	1.55	224	−0.41	.684	−0.06	[−0.32, 0.21]
Tension	3.70	1.31	3.59	1.34	224	0.61	.542	0.08	[−0.18, 0.34]
Extrinsic	4.62	1.42	4.48	1.27	223.24	0.79	.430	0.11	[−0.16, 0.37]

Note. Modified *df*, *t* statistic, and *p* values are presented when Levene's Test for Equality of Variance was violated.

^a Without these exclusions, the simple effect of removing pay on performance remained significant, as did the main effect of type of effect on performance and number attempted. No other effects on the dependent measures emerged.

conduct simple-effects tests of reward within type of effect condition (see Table 5). The hypothesized classic undermining effect of pay on number attempted was marginally significant. Number attempted in the free-choice period was marginally lower in the pay removed than in the pay never offered condition. Number attempted in the free-choice phase in the gamification removed and the gamification never offered conditions did not differ.

Performance was calculated as the number of weighted correct n-back decisions (one point for 1-back, two points for 2-back, and three points for 3-back) during the free-choice phase. Performance scores were submitted to a 2 (Type of effect: undermining pay vs. undermining gamification) $\times$ 2 (Reward offered: never offered vs. removed) between-participants ANOVA (see Table 3). There were again large effects of undermining type, such that participants

Table 5*Independent Samples t-Tests of Reward, by Type of Effect*

Variable	Pay never offered		Pay removed		df	t	p	d	95% CI _d
	n = 108		n = 116						
	M	SD	M	SD					
Attempts	2.95	3.58	2.10	3.10	212.16 ^a	1.89	.060	0.25	[−0.01, 0.52]
Performance	3.00	4.87	1.54	3.08	178.43 ^a	2.66	.009	0.36	[0.09, 0.62]
	Gamification never offered		Gamification removed		df	t	p	d	95% CI _d
	n = 107		n = 110						
	M	SD	M	SD					
Attempts	5.66	3.51	5.88	3.17	211.47	0.48	.631	−0.07	[−0.33, 0.20]
Performance	7.64	6.60	7.52	6.15	215	0.15	.884	0.02	[−0.25, 0.29]

Note. Modified *df*, *t* statistic, and *p* values are presented when Levene's Test for Equality of Variance was violated.

^a Without these exclusions, the simple effect of removing pay on performance remained significant, as did the main effect of type of effect on performance and number attempted. No other effects on the dependent measures emerged.

performed better in the undermining gamification conditions (when performance-based pay was always provided) than in the undermining pay conditions (when gamification was always provided). There was no overall effect of reward offered, and the predicted interaction was not significant. We again conducted simple-effects tests within type of effect (see Table 5). The hypothesized classic undermining effect of pay was in the predicted direction and significant. Free-choice performance was better in the pay never offered condition than in the pay removed condition. However, free-choice performance in the gamification removed and gamification never offered conditions did not significantly differ.

Discussion

We deviated slightly from the preregistered analysis plan to exclude participants who did not pass the test of understanding the instructions, though doing so did not change our findings. We observed a small undermining effect of removing pay for performance, and a marginal effect on number attempted during a free choice period. No significant undermining effects were observed on gamification. However, the predicted interaction did not emerge on these measures. Thus, undermining effects of gamification were not significantly smaller than undermining effects of pay.

In addition to the issue of some participants failing to understand the instructions, our measures of extrinsic motivation failed to reveal the predicted effects. This could be due to the performance-contingent pay being too small to be experienced as controlling. Past work suggests that performance-contingent bonuses are quite impactful on work productivity, even when they are as small as 3% of base pay (Bucklin & Dickinson, 2001). The bonus pay provided in Study 1 was up to 45% of base pay, contingent on correct location decisions. Indeed, we observed large beneficial effects of offering performance-contingent pay for task engagement and performance. Instead, we believe the lack of effects on self-reports of extrinsic motivation resulted from participants confusing the pay to participate in the study with the performance-based bonus pay. Finally, it seems that undermining effects in this online paradigm are smaller than in prior studies, meaning that our study was underpowered to detect the predicted interaction. We conducted a second study to address these limitations.

Study 2

Study 2 was a close replication of Study 1. We reworded items related to extrinsic motivation and instrumentality of making money to specify earning a bonus based on performance. We also added a test of understanding the instructions to the free-choice phase of the study to ensure that participants understood any changes in the task. We then preregistered an analysis plan that called for excluding data from participants that failed either test of understanding in addition to the same exclusions that were planned for Study 1. Finally, given the smaller than expected effects sizes observed in Study 1, we conducted a power analysis based on the effects observed in that study. Our hypotheses were the same as those posed in Study 1 and were also preregistered (see Preregistered Material).

Method

Participants and Design

Participants were recruited from MTurk in return for \$0.30 payment and the possibility of a bonus for performance. Participants had to be located in the U.S. and have a 95% acceptance rate on a minimum of 100 previous HITs. They were also told prior to starting the study that they would need a number-pad on their keyboard. Participant recruitment was conducted with CloudResearch Turk-Prime (Litman et al., 2017), which additionally blocked anyone who previously completed the study or shared an IP address with a previous participant.

Power simulations (Lakens & Caldwell, 2020) suggested that a sample size of $N = 1,200$ would be sufficient to detect a significant interaction with 80% power, assuming that the pattern of means observed on number attempted during the free-choice phase in Study 1. We, therefore, preregistered collecting data until $N = 1,200$ had completed the study.

As in Study 1, participants were randomly assigned to one condition in a 2 (Type of effect: undermining pay vs. undermining gamification) $\times$ 2 (Reward offered: never offered vs. removed) between-participants design. Data analysis did not proceed until data collection was completed.

Location Memory Task

The location memory tasks were the same as in Study 1.

Manipulation Checks

The effectiveness of the manipulations in the initial phase was assessed with self-report measures adapted from Study 1. Items regarding extrinsic motivation were reworded to concern the bonus payment (e.g., “I am only interested in playing this memory game for a bonus based on performance”). Likewise, the instrumentality measure was altered, such that the money goal referred to “earning bonus payment.”¹ Participants again provided their age and gender in open-ended responses.

Procedure

The study proceeded as in Study 1, with the exception that a second test of understanding was administered after the instructions for the free-choice phase. Specifically, participants were asked: “I should remember the locations of the (gopher, circle, tree, horse)”; Remembering the correct location is rewarded with a monetary bonus (True or False),” and “I can quit the task at any time by pressing the escape key (True or False).” They received feedback on the correct answers and the free-choice phase began.

Results

Data Preparation

All analyses were conducted with R (R Core Team, 2020) using the Remdr package (Fox, 2005). The sample consisted of $N = 1,298$ complete responses. Additional 1,130 incomplete responses were

¹ Due to a programming oversight, the “earning money” label was still used in the gamification never offered condition.

recorded. Incompletes were slightly higher in the gamification never offered condition (52% of recorded responses incomplete) compared to the initially gamified conditions (45% of recorded responses incomplete), $\chi^2(3) = 9.99$, $p = .019$. We suspect that this difference would work against our hypotheses if individuals were discontinuing the study due to boredom, though roughly half of incomplete responses involved participants who did not proceed beyond the initial instructions.

As preregistered, we excluded 54 participants who did not respond to any trial of the memory task. These exclusions did not differ by condition ($p = .73$). We excluded additional 139 participants for getting 50% or less on either test of understanding the instructions. These exclusions were slightly higher in the pay never offered condition (5% overall) compared to other conditions (3.4% to 4.1% overall), $F(31,294) = 3.18$, $p = .023$. No additional exclusions were made for failed attention checks or identical responding. The final sample size was, therefore, $N = 1,105$ (Age: $M = 40.84$, $SD = 13.18$; Gender: Male = 436, Female = 652, Other = 11, Did Not Identify = 6), with all conditions greater than $n = 250$. This sample size provides 80% power to detect a small-to-moderate undermining effect, as we observed in Study 1 ($ds > .24$). Among retained responses, scores on the test of understanding the instructions were high ($M_{\text{part1}} = 88.3\%$; $M_{\text{part2}} = 89.1\%$). Correlations, grand means, and standard deviation of all study variables are presented in Table 6. Controlling for age or gender of participants did not alter the results, so we do not discuss these variables further.

Manipulation Checks

The manipulation check measures assessed at the end of the initial task phase were submitted to 2 (Type of effect: undermining pay vs. undermining gamification) $\times$ 2 (Reward offered: never offered vs. removed) between-participants ANOVAs (see Table 7). As expected, no significant differences were found for effort and instrumentality of improving memory. Means on both measures were high, indicating that individuals were generally engaged by the task and believed that the task was valid.

The predicted interactions were all significant. Simple effects of type were conducted within reward offered condition (see Table 8). Those in the pay never offered (gamified) condition reported higher

instrumentality for having fun and less instrumentality for earning bonus payment, compared to individuals in the gamification never offered (pay) condition. No significant differences were observed in the removal condition. Similarly, those in the pay never offered condition reported more enjoyment, less extrinsic motivation, higher competence, and less tension, compared to those in the gamification never offered condition. In contrast, there were no significant effects in the removal condition.

In sum, the manipulations were successful. Participants reported enjoying the gamified versions of the task (at or above the scale midpoints on enjoyment on instrumentality for fun) and were sufficiently extrinsically motivated by the presence of monetary bonuses (at or above the scale midpoints on extrinsic motivation and instrumentality for earning bonus money).

Hypothesis Tests on Free Choice Behavior

The number of blocks attempted during the free-choice phase was submitted to a 2 (Type of effect: undermining pay vs. undermining gamification) $\times$ 2 (Reward offered: never offered vs. removed) between-participants ANOVA (see Table 7). As in Study 1, there were large effects of type of undermining effect, such that participants attempted far more items in the undermining gamification conditions (in which performance-contingent pay was always provided) than in the undermining pay conditions (in which gamification was always provided). Providing performance-contingent pay thus had large benefits for free-choice engagement.

There was also a significant main effect of reward offered on number attempted, such that removing a reward led to decreased task engagement during the free-choice phase, compared to never offering the reward. Examining each type of undermining effect independently revealed a significant undermining effect of pay. The undermining effect of gamification was in the same direction, but was not significant. The predicted interaction was not significant, indicating that the size of the undermining effects did not significantly differ.

Performance was scored in the same manner as in Study 1. Submitting these scores to a 2 (Type of effect: undermining pay vs. undermining gamification) $\times$ 2 (Reward offered: never offered vs. removed) between-participants ANOVA (see Table 7) revealed very similar effects as observed on number attempted. The predicted

Table 6
Correlations for Study 2 Variables

Variable	1	2	3	4	5	6	7	8	9	10
1. Memory	—									
2. Fun	.430	—								
3. Money	.240	.151	—							
4. Enjoy	.271	.582	.080	—						
5. Effort	.175	.134	.121	.340	—					
6. Competence	.181	.227	.093	.396	.005	—				
7. Tension	-.017	-.185	.027	-.200	.158	-.411	—			
8. Extrinsic	-.099	-.316	.264	-.372	-.052	-.047	.153	—		
9. Attempts	.050	.040	.038	.117	.082	.037	.070	-.042	—	
10. Performance	.025	-.014	.059	.066	.049	.066	.062	.032	.886	—
<i>M</i>	4.86	4.31	4.88	5.06	5.94	3.67	3.66	3.85	4.60	6.29
<i>SD</i>	1.83	2.02	1.97	1.45	0.97	1.45	1.36	1.49	3.70	6.35

Table 7*ANOVA Results for Manipulation Checks, Trial Attempts, and Performance (Study 2)*

Variable	Reward offered				Type of effect				Interaction			
	<i>F</i>	<i>p</i>	<i>d</i>	95% CI _d	<i>F</i>	<i>p</i>	<i>d</i>	95% CI _d	<i>F</i>	<i>p</i>	<i>d</i>	95% CI _d
Memory	0.00	.961	.00	[−.12, .12]	1.19	.276	.07	[−.05, .18]	0.21	.647	.03	[−.09, .15]
Fun	11.70	<.001	.21	[.09, .32]	15.26	<.001	.24	[.12, .35]	24.16	<.001	.30	[.18, .41]
Money	8.94	.003	.18	[.06, .30]	2.61	.107	.10	[−.02, .22]	7.97	.005	.17	[.05, .29]
Enjoy	15.65	<.001	.24	[.12, .36]	8.65	.003	.18	[.06, .30]	27.80	<.001	.32	[.20, .44]
Effort	0.09	.762	.02	[−.10, .14]	0.61	.436	.05	[−.07, .17]	0.02	.877	.01	[−.10, .13]
Competence	0.00	.985	.00	[−.12, .12]	6.11	.014	.15	[.03, .27]	10.04	.002	.19	[.07, .31]
Tension	0.08	.772	.02	[−.10, .14]	5.58	.018	.14	[.02, .26]	8.49	.004	.18	[.06, .29]
Extrinsic	0.01	.933	.01	[−.11, .12]	13.02	<.001	.22	[.10, .34]	25.53	<.001	.30	[.19, .42]
Attempts	4.22	.040	.12	[.01, .24]	266.24	<.001	.98	[.86, 1.11]	0.80	.372	.05	[−.06, .17]
Performance	8.12	.004	.17	[.05, .29]	303.79	<.001	1.05	[.92, 1.18]	0.20	.655	.03	[−.09, .15]

Note. $df_1 = 1$, $df_2 = 1,101$.

interaction was not significant, but there was again an overall effect of reward offered. Removing gamification significantly reduced performance relative to never offering gamification. Removing pay marginally reduced performance relative to never offering pay (see Table 9).

Exploratory Analyses

There was an effect of reward offered on performance in the initial phase, $F(1, 1,101) = 4.66$, $p = .031$, such that scores were higher in the reward never offered conditions. There were no effects of type,

$F_s < 1.25$, $p_s > .265$. Controlling for initial phase performance slightly reduced the effect of reward offered on number attempted in the free choice phase ($p = .069$), but the effect of reward on performance in the free choice phase remained significant ($p = .017$). One-way ANOVAs revealed condition differences on both tests of understanding the instructions ($p_s < .01$). When controlling for these scores, the main effect of reward on performance remained significant ($p = .007$), whereas the effect on number attempted was slightly reduced ($p = .061$).

We also conducted exploratory mediation analyses using the R mediation package (Tingley et al., 2014), focusing on the effects of

Table 8*Independent Samples *t*-Tests of Type of Effect, by Reward Condition (Study 2)*

Variable	Pay never offered		Gamification never offered		<i>df</i>	<i>t</i>	<i>p</i>	<i>d</i>	95% CI _d
	<i>n</i> = 268		<i>n</i> = 251						
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>					
Memory	4.95	1.79	4.78	1.89	517	1.06	.291	.09	[−.08, .27]
Fun	4.61	1.93	3.55	2.03	517	6.06	<.001	.53	[.36, .71]
Money	4.44	2.14	4.96	2.00	517	2.88	.004	.25	[.08, .43]
Enjoy	5.23	1.35	4.52	1.56	494.9	5.48	<.001	.48	[.31, .66]
Effort	5.96	0.96	5.90	0.98	517	0.64	.521	.06	[−.12, .23]
Competence	3.91	1.44	3.42	1.41	517	3.91	<.001	.34	[.17, .52]
Tension	3.46	1.33	3.89	1.42	517	3.57	<.001	.31	[.14, .49]
Extrinsic	3.47	1.53	4.24	1.56	517	5.68	<.001	.50	[.32, .67]
	Pay removed		Gamification removed		<i>df</i>	<i>t</i>	<i>p</i>	<i>d</i>	95% CI _d
	<i>n</i> = 277		<i>n</i> = 309						
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>					
Memory	4.89	1.86	4.82	1.79	584	0.46	.644	.04	[−.12, .20]
Fun	4.43	2.05	4.55	1.91	565.6	0.73	.464	.06	[−.10, .22]
Money	5.13	1.89	4.98	1.80	584	0.93	.351	.08	[−.09, .24]
Enjoy	5.11	1.49	5.31	1.29	548.5	1.73	.085	.14	[−.02, .31]
Effort	5.97	0.95	5.93	0.99	584	0.45	.649	.04	[−.12, .20]
Competence	3.64	1.48	3.70	1.42	584	0.50	.615	.04	[−.12, .20]
Tension	3.67	1.33	3.63	1.33	584	0.41	.683	.03	[−.13, .20]
Extrinsic	3.93	1.47	3.80	1.34	561.7	1.10	.273	.09	[−.07, .25]

Note. Modified *df*, *t* statistic, and *p* values are presented when Levene's Test for Equality of Variance was violated.

Table 9*Independent Samples t-Tests of Reward, by Type of Effect (Study 2)*

Variable	Pay never offered <i>n</i> = 268		Pay removed <i>n</i> = 277		<i>df</i>	<i>t</i>	<i>p</i>	<i>d</i>	95% CI _d
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>					
Attempts	3.25	3.59	2.66	3.36	537.8	1.98	.049	.17	[.00, .34]
Performance	3.74	5.22	2.92	4.72	533.5	1.91	.057	.16	[.00, .33]
	Gamification never offered <i>n</i> = 251		Gamification removed <i>n</i> = 309		<i>df</i>	<i>t</i>	<i>p</i>	<i>d</i>	95% CI _d
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>					
Attempts	6.34	3.29	6.11	3.03	558	0.87	.387	.07	[−.09, .24]
Performance	9.80	6.39	8.68	6.02	558	2.13	.034	.18	[.01, .35]

Note. Modified *df*, *t* statistic, and *p* values are presented when Levene's Test for Equality of Variance was violated.

extrinsic motivation and enjoyment as most theoretically relevant. We defined a model in which the interaction of reward offered and type of effect predicted extrinsic motivation and enjoyment (a paths), and the interaction of each mediator and type of effect predicted the dependent measure (b paths). Reward offered (never offered = −1; removed = 1) and type of effect (undermining pay = −1; undermining gamification = 1) were effect coded. Indirect effects were estimated at each level of type of effect. These analyses were conducted separately for number attempted and performance in the free-choice phase (see Table 10).

According to SDT (Deci et al., 1999; Ryan & Deci, 2000), we would expect the undermining effect of pay to be due to reduced enjoyment when pay is introduced. In our analyses, we instead observed a significant negative indirect effect through extrinsic motivation. Extrinsic motivation was higher when pay was present in the initial phase than when it was never offered. However, extrinsic motivation predicted fewer items attempted in the

free-choice phase. That is, extrinsic motivation had a negative association with task engagement in the absence of pay. There was no indirect effect through enjoyment. Enjoyment did not decrease when pay was added nor did enjoyment significantly predict number attempted.

If game features are experienced in the same way as other performance-contingent rewards, we would again expect the undermining effect of gamification to involve reduced enjoyment or a reduced association between enjoyment and free-choice behavior. For the undermining effect of removing gamification on number attempted, there was a significant positive indirect effect through enjoyment. Enjoyment was higher when gamification was offered in the initial phase compared to when gamification was never offered. However, enjoyment predicted higher engagement even when gamification was not present in the free-choice phase. These results are counter to the undermining effect, consistent with a suppression effect. Undermining effects on number attempted could, therefore,

Table 10*Mediation Models*

Effects	Extrinsic			Intrinsic			Number attempted			Performance		
	<i>b</i>	<i>t</i>	<i>p</i>	<i>b</i>	<i>t</i>	<i>p</i>	<i>b</i>	<i>t</i>	<i>p</i>	<i>b</i>	<i>t</i>	<i>p</i>
Reward offered	.004	0.08	.934	.170	3.96	<.001	−.221	2.22	.027	−.524	3.07	.002
Type of effect	.160	3.61	<.001	−.126	2.94	.003	−.904	1.68	.092	−1.290	1.40	.161
Reward offered × Type of effect	−.224	5.05	<.001	.226	5.27	<.001						
Extrinsic							−.070	0.99	.321	.148	1.23	.218
Intrinsic							.368	5.08	<.001	.527	4.24	<.001
Extrinsic × Type of effect							.411	5.83	<.001	.610	5.05	<.001
Intrinsic × Type of effect							.200	2.76	.006	.381	3.07	.002
Conditional indirect effects	<i>b</i>			95% CI			<i>b</i>			95% CI		
Removing pay												
Extrinsic			−.109			[−.194, −.050]			−.105			[−.187, −.033]
Intrinsic			−.010			[−.051, .008]			−.008			[−.045, .017]
Removing gamification												
Extrinsic			−.075			[−.151, −.030]			−.167			[−.313, −.061]
Intrinsic			.225			[.132, .363]			.359			[.197, .554]

Note. *N* = 1,105. Conditional indirect effects bootstrapped with 500 simulations.

not be explained by reduced enjoyment. There was a significant negative indirect effect through extrinsic motivation. Extrinsic motivation was higher when gamification was never offered compared to when gamification was offered. Extrinsic motivation then predicted higher engagement when gamification was not present in the free-choice phase. Thus, gamification in the initial phase lowered extrinsic motivation, which in turn was associated with lower task engagement in the free-choice phase.

Similar effects were observed on performance. Extrinsic motivation statistically mediated the undermining effect of removing pay on performance, and there was no indirect effect through enjoyment. The undermining effect of removing gamification was increased when controlling for enjoyment, and was statistically mediated by extrinsic motivation.

Parallel analyses with the instrumentality of fun and of earning bonus money as the candidate mediators did not reveal any significant indirect effects. In sum, extrinsic motivation mediated both types of undermining effect, though in slightly different ways. The undermining effect of pay was mediated through a negative association between extrinsic motivation and free-choice behavior. The undermining effect of gamification was instead mediated by a reduction in extrinsic motivation when introducing gamification.

General Discussion

The present work was designed to replicate and extend past research on the undermining effects of removing reward. First, we attempted to replicate the classic undermining effects of pay (e.g., Deci, 1971) in an online work environment using a gamified memory task. Second, we sought to test whether undermining effects would also be observed when removing game features of the task. This undermining effect of gamification could emerge either, because certain game features are experienced as extrinsically motivating or because removing any type of previously available reward is demotivating. Consistent with SDT explanations of the undermining effect (Deci et al., 1999; Ryan & Deci, 2000), we predicted an interaction, such that we would only observe undermining effects of performance-based pay. We preregistered our hypotheses and analysis plans, deviating only slightly from the latter in Study 1.

Contrary to expectation and despite a large sample size, we did not observe a significant interaction in either study. In Study 1, we observed a significant undermining effect of pay on performance and a marginal effect on number of blocks attempted. The effects on gamification were not significant. With a much larger sample size in Study 2, we observed a significant overall undermining effect of removing reward on both number attempted and performance. In fact, removing gamification had a significant undermining effect on performance. Given the somewhat disparate results across studies, we conducted a meta-analysis combining the main effects of reward offered in Studies 1 and 2 (Lipsey & Wilson, 2001). Reward offered was significant for both number of blocks attempted, $d = .11$, 95% CI = [.01, .21], and performance, $d = .16$, 95% CI = [.06, .26].² Within type of effect, removing pay had robust effects on both number attempted, $d = .19$, 95% CI = [.05, .34], and performance, $d = .22$, 95% CI = [.08, .36]. Removing gamification had slightly smaller effects on number attempted, $d = .03$, 95% CI = [−.11, .17], and performance, $d = .13$, 95% CI = [−.01, .28], with the latter approaching significance. These findings suggest that removing

either type of reward can undermine free-choice behavior. Our estimation of the effects of removing performance-contingent pay was slightly smaller than (but well within the confidence interval of) prior meta-analyses of the undermining effect (Cameron et al., 2001; Cameron & Pierce, 1994; Deci et al., 1999, 2001; Eisenberger et al., 1999; Eisenberger & Cameron, 1996; Rummel & Feinberg, 1988; Tang & Hall, 1995; Wiersma, 1992). Despite our attempts to ensure that participants were paying attention, understood the instructions, and were motivated by the incentive of bonus payment, we did not observe large undermining effects. It is possible that our effects were smaller due to the type of task selected, reliance on an online sample that participates for pay (rather than undergraduate participants or children as in prior studies), or bonus pay that was still too small to undermine intrinsic interest. It is also possible that paradigms with larger effect sizes would be able to detect differences between the undermining effects of pay and gamification. We discuss these points in more detail below.

Despite finding broader undermining effects, exploratory mediation analyses suggested that the exact mechanisms behind these effects may not be the same. For the classic undermining effect of removing pay, our analyses revealed indirect effects through extrinsic motivation. Specifically, individuals who were extrinsically motivated when pay was presented in the initial phase were more likely to abandon the task early when pay was not offered. There were no indirect effects through intrinsic motivation.

A different pattern emerged for the undermining effects of gamification. Both enjoyment and extrinsic motivation continued to predict higher task engagement and performance when gamification was not offered in the free-choice phase. Rather, it seems that adding gamification reduced extrinsic motivation during the initial phase. Importantly, it does not seem that gamification is necessarily experienced as controlling, as participants reported experiencing more competence and less tension on the gamified task.

Finally, there was a large and reliable benefit to engagement and performance when offering performance-based pay during the task, relative to gamification. Across both studies, individuals attempted a greater number of items and performed better when offered performance-based monetary rewards, compared to when the task was only gamified.

Implications for Theoretical Explanations of Undermining Effects

Overall, our results provide evidence supportive of reduced task engagement and performance when removing incentives. However, our results do not comport perfectly with any one theoretical perspective. Although we replicated the classic undermining effect as predicted by SDT (Deci et al., 1999; Ryan & Deci, 2000), we did not observe that there is a unique effect of removing extrinsic rewards. Gamification had nonsignificantly smaller undermining effects relative to performance-contingent pay. In addition, exploratory analyses suggested that the undermining effect of pay was statistically mediated through a negative association with extrinsic motivation, rather than reduced enjoyment. The undermining effects of removing gamification were not driven by “controlling” extrinsic motivation. In fact, adding game features reduced extrinsic

² All effects remained significant when including participants who failed the test of understanding the instructions in Study 1.

motivation, which in turn predicted lower task engagement and performance. On the other hand, increased enjoyment may provide a buffer against these effects. Finally, the benefits of performance-based pay on task engagement and performance were considerably larger than the undermining effects that we observed.

As an alternative perspective, we discussed the Structural Model of Intrinsic Motivation (Kruglanski et al., 2018) and the notion that added incentives can crowd-out the effects of existing incentives. Consistent with this perspective, we observed undermining effects of both monetary reward and gamification. Gamification in particular seemed to reduce the benefits of pay for extrinsic motivation. However, we did not see parallel effects of pay reducing the benefits of enjoyment, as would be expected by a “crowding out” effect. Although our manipulations affected the instrumentality measures as expected, these measures did not predict free-choice behavior. In sum, there was not straightforward evidence in support of this theoretical explanation either.

Given these findings, future studies will need to examine the underlying process behind the consequences of removing incentives. Theoretical accounts may need to be revised to consider that different types of incentives may operate through different mechanisms to influence free-choice behavior. It is also possible that the mechanisms behind undermining effects depend in part on the task context. As discussed below, we recruited a sample of workers who are likely to be extrinsically motivated. This could have increased the role of extrinsic motivation in our effects.

Implications for Application

The overall pattern of results suggests that the presentation of tangible rewards for performance may not be so ill-advised, at least for tasks of the nature utilized here and particularly when these rewards can be continued throughout the task. We observed only small undermining effects of pay that were overwhelmed by the very large benefits to task engagement and performance of providing small monetary awards throughout the task. These findings are consistent with past work in applied contexts, demonstrating that extrinsic reward can be beneficial under the right circumstances (Cerasoli et al., 2014; Eisenberger & Cameron, 1996).

Likewise, we observed that efforts to make tasks more enjoyable can undermine task engagement and performance. In particular, the provision of points and feedback for completing task-relevant actions lead to small undermining effects (see also Amriani et al., 2013; Hanus & Fox, 2015). The transition from a “fun” task to a more serious task may, therefore, be detrimental to persistence and performance. On the other hand, we observed that gamification increased enjoyment, which was associated with higher persistence and performance. Indeed, it is impressive that merely introducing a fun version of the task resulted in some amount of task persistence in a group of participants that had no other reason to continue participating. In fact, continuing to work on the task represented an opportunity cost for MTurk participants who could have easily moved on to new paid work. From this point of view, it would not have been particularly surprising if we had observed zero task engagement in the conditions with pay removed or pay never offered. If any amount of task engagement can be achieved with less expense, this might be of interest for applications in online tasks in which compensation is not practically feasible. These findings are consistent with past work on the benefits of gamification

(Domínguez et al., 2013). We also observed benefits of gamification in terms of increased enjoyment and competence, and reduced tension, supporting the view that gamification increases intrinsic interest (Wouters et al., 2013). Specifically, introducing cartoon-like targets, providing points, and giving participants positive visual feedback for tracking the location of the targets seemed sufficient to make the task more enjoyable. This supports the popular usage of these game features (Dicheva et al., 2015), as well as underlying research demonstrating that specific and positive feedback can increase self-efficacy and task enjoyment (Bandura & Locke, 2003; Harackiewicz, 1979; Harackiewicz et al., 1984).

Limitations and Generalizability

Examining the manipulation checks, it appears that we were successful in creating an online task that was enjoyable and thus represented a valid test of the undermining hypothesis. The “whack-a-mole” version of the task was rated as more enjoyable and better served the goal of having fun, compared to the paid version with the colored circles. Moreover, this version of the task resulted in less tension and more feelings of competence, though these effects were only small to moderate in size.

The task manipulation had no effect on measures of extrinsic motivation in Study 1. However, more explicitly referring to the availability of performance-contingent bonus payment seemed to alleviate this issue. Nonetheless, extrinsic motivation was rated as near the midpoint even when pay was never offered. This is surprising, given that the presentation of tangible rewards had very large effects on task engagement and performance. Future studies should consider wording these questions to explicitly exclude compensation unrelated to engaging in the task itself. It is also possible that recruiting MTurk workers results in a sample that is high in extrinsic motivation. Specifically, MTurk workers may be more extrinsically motivated than groups of volunteers or participants with a more inherent interest in the task (see also Hagger & Chatzisarantis, 2011).

A related point is that, even in the gamification never offered condition, the task was rated as somewhat enjoyable. It is possible that the location memory task was relatively lively compared to most MTurk tasks. Thus, removing the game features may not have undermined intrinsic motivation as much as it would seem. Future research on gamification could, therefore, attempt to improve upon the present study by making the control task even less interesting.

We powered both studies to detect small-to-moderate effects, but our observed effects tended to be smaller than expected. Future studies of the undermining effect in online samples should likely be powered to detect even smaller effects. In addition, we aimed to achieve 80% power, but future studies may aim to achieve 90%–95% power given the possibility of a small effect (Curran-Everett, 2017). Given these considerations, it may be necessary to examine situations in which larger effects could emerge or to develop within-participant designs.

Some final limitations of the present study involve generalizability. It is likely that task demands impact intrinsic interest. Existing research on the undermining effect has determined that tasks emphasizing quality of performance, those more complex and that require more personal cognitive investment, are more intrinsically motivating (Cerasoli et al., 2014). This is compared to tasks that focus on quantity of performance (produced by focused,

persistent, and structured behavior; Gilliland & Landis, 1992). Though not particularly complex, the n-back task used in the present study fits criteria for emphasizing quality, because it demands higher levels of cognitive effort (Jaeggi et al., 2010). Undermining effects of pay are typically only observed for tasks that are initially intrinsically interesting (Cameron et al., 2001; Deci et al., 1999), and thus, our effects likely would not be the same on tasks more heavily emphasizing quantity of performance. Additional studies comparing undermining effects of pay and gamification on more complex tasks and on tasks emphasizing quantity of performance are, therefore, needed. Tasks emphasizing quality of performance are also important to further research, as they are particularly relevant to work, educational, and recreational contexts.

Relatedly, we used relatively small monetary rewards and a short task duration. Many applications of gamification, such as serious games, involve longer durations of time (e.g., educational applications that require multiple sessions) and, therefore, involve larger rewards. Even for short tasks, larger rewards may produce stronger extrinsic motivation and thus stronger undermining effects. However, we note that even small bonuses relative to base payment seem to be effective in motivating behavior in work contexts (Bucklin & Dickinson, 2001). It is likely that undermining effects are more dependent on performance-contingency, as found in prior meta-analyses (e.g., Cameron et al., 2001; Deci et al., 1999). These limitations should be addressed in future studies.

Finally, we focused on one aspect of gamification involving the provision of points for task-relevant actions. Future work should examine whether other types of gamification, including cooperation and competition, complex narratives, or freedom of choice for the player, have undermining effects. In theory, such game aspects should effectively increase intrinsic interest (Mekler et al., 2017). On the other hand, our findings suggest that reducing extrinsic motivation can also be associated with reduced task engagement when these features are subsequently removed.

Conclusion

This study advances literature on the undermining effect in several ways. First, we found evidence for a small undermining effect on free-choice behavior with a large sample and pre-registered analysis. Undermining effects of pay were not always statistically significant, but of a size consistent with prior meta-analyses of the literature. These results suggest that the undermining effect extends to tasks similar to those common in online “brain-training” programs (e.g., BrainHQ, Luminosity), and with a sample of online workers given an explicit improvement goal. Undermining effects of gamification were also observed, and were not significantly smaller than the undermining effects of pay. This suggests that gamification may not completely solve the issue of how to maintain task interest. Contrary to theory, undermining effects were mediated by extrinsic motivation rather than reduced enjoyment. Future research will be needed to replicate and extend these findings as well as further develop theoretical explanations of the undermining effect. Finally, despite the potential drawbacks of removing incentives, the evidence suggests that (at least with short tasks and for a sample of paid participants) the benefits of providing small performance-based tangible rewards are considerable. Gamification is a cost-effective way to promote task engagement, but may not be as effective.

References

- Amirani, A., Aji, A. F., Utomo, A. Y., & Junus, K. M. (2013). An empirical study of gamification impact on e-Learning environment. *Proceedings of 2013 3rd International Conference on Computer Science and Network Technology* (pp. 265–269).
- Arkes, H. R. (1979). Competence and the overjustification effect. *Motivation and Emotion*, 3(2), 143–150. <https://doi.org/10.1007/BF01650599>
- Aron, A., Aron, E. A., & Smollan, D. (1992). Inclusion of other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, 63(4), 596–612. <https://doi.org/10.1037/0022-3514.63.4.596>
- Bandura, A., & Locke, E. A. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology*, 88(1), 87–99. <https://doi.org/10.1037/0021-9010.88.1.87>
- Bucklin, B. R., & Dickinson, A. M. (2001). Individual Monetary Incentives. *Journal of Organizational Behavior Management*, 21(3), 45–137. https://doi.org/10.1300/J075v21n03_03
- Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14(5), 365–376. <https://doi.org/10.1038/nrn3475>
- Cameron, J., Banko, K. M., & Pierce, W. D. (2001). Pervasive negative effects of rewards on intrinsic motivation: The myth continues. *The Behavior Analyst*, 24(1), 1–44. <https://doi.org/10.1007/BF03392017>
- Cameron, J., & Pierce, W. D. (1994). Reinforcement, reward, and intrinsic motivation: A meta-analysis. *Review of Educational Research*, 64(3), 363–423. <https://doi.org/10.3102/00346543064003363>
- Cameron, J., Pierce, W. D., Banko, K. M., & Gear, A. (2005). Achievement-based rewards and intrinsic motivation: A test of cognitive mediators. *Journal of Educational Psychology*, 97(4), 641–655. <https://doi.org/10.1037/0022-0663.97.4.641>
- Cerasoli, C. P., Nicklin, J. M., & Ford, M. T. (2014). Intrinsic motivation and extrinsic incentives jointly predict performance: A 40-year meta-analysis. *Psychological Bulletin*, 140(4), 980–1008. <https://doi.org/10.1037/a0035661>
- Cerasoli, C. P., Nicklin, J. M., & Nassrelgragawi, A. S. (2016). Performance, incentives, and needs for autonomy, competence, and relatedness: A meta-analysis. *Motivation and Emotion*, 40(6), 781–813. <https://doi.org/10.1007/s11031-016-9578-2>
- Curran-Everett, D. (2017). CORP: Minimizing the chances of false positives and false negatives. *Journal of Applied Physiology*, 122(1), 91–95. <https://doi.org/10.1152/jappphysiol.00937.2016>
- de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, 47(1), 1–12. <https://doi.org/10.3758/s13428-014-0458-y>
- Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, 18(1), 105–115. <https://doi.org/10.1037/h0030644>
- Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125(6), 627–668. <https://doi.org/10.1037/0033-2909.125.6.627>
- Deci, E. L., Koestner, R., & Ryan, R. M. (2001). Extrinsic rewards and intrinsic motivation in education: Reconsidered once again. *Review of Educational Research*, 71(1), 1–27. <https://doi.org/10.3102/00346543071001001>
- Deci, E. L., & Ryan, R. M. (1985). Cognitive evaluation theory. *Intrinsic motivation and self-determination in human behavior Perspectives in Social Psychology* (pp. 43–85). Springer. https://doi.org/10.1007/978-1-4899-2271-7_3
- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: defining “gamification”. *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments* (pp. 9–15). Association for Computing Machinery.

- Dicheva, D., Dichev, C., Agre, G., & Angelova, G. (2015). Gamification in education: A systematic mapping study. *Educational Technology & Society*, 18(3), 75–88.
- Domínguez, A., Saenz-de-Navarrete, J., de-Marcos, L., Fernández-Sanz, L., Pagés, C., & Martínez-Herráiz, J.-J. (2013). Gamifying learning experiences: Practical implications and outcomes. *Computers & Education*, 63, 380–392. <https://doi.org/10.1016/j.compedu.2012.12.020>
- Eisenberger, R., & Cameron, J. (1996). Detrimental effects of reward. Reality or myth? *American Psychologist*, 51(11), 1153–1166. <https://doi.org/10.1037/0003-066X.51.11.1153>
- Eisenberger, R., Pierce, W. D., & Cameron, J. (1999). Effects of reward on intrinsic motivation: Negative, neutral, and positive: Comment on Deci, Koestner, and Ryan (1999). *Psychological Bulletin*, 125(6):677–691. <https://doi.org/10.1037/0033-2909.125.6.677>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Ferrara, J. (2013). Games for persuasion: Argumentation, procedurality, and the lie of gamification. *Games and Culture*, 8(4), 289–304. <https://doi.org/10.1177/1555412013496891>
- Fox, J. (2005). Getting started with the R commander: A basic-statistics graphical user interface to R. *Journal of Statistical Software*, 14(9), 1–42. <https://doi.org/10.18637/jss.v014.i09>
- Gilliland, S. W., & Landis, R. S. (1992). Quality and quantity goals in a complex decision task: Strategies and outcomes. *Journal of Applied Psychology*, 77(5), 672–681. <https://doi.org/10.1037/0021-9010.77.5.672>
- Giner-Sorolla, R. (2018). *Powering your interaction. Approaching Significance*. <https://approachingblog.wordpress.com/2018/01/24/powering-your-interaction-2/>
- Green, A., Ellis, K. A., Ellis, J., Bartholomeusz, C. F., Ilic, S., Croft, R. J., Phan, K. L., & Nathan, P. J. (2005). Muscarinic and nicotinic receptor modulation of object and spatial n-back working memory in humans. *Pharmacology, Biochemistry, and Behavior*, 81(3), 575–584. <https://doi.org/10.1016/j.pbb.2005.04.010>
- Hagger, M. S., & Chatzisarantis, N. L. (2011). Causality orientations moderate the undermining effect of rewards on intrinsic motivation. *Journal of Experimental Social Psychology*, 47(2), 485–489. <https://doi.org/10.1016/j.jesp.2010.10.010>
- Hanus, M. D., & Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers & Education*, 80, 152–161. <https://doi.org/10.1016/j.compedu.2014.08.019>
- Harackiewicz, J. M. (1979). The effects of reward contingency and performance feedback on intrinsic motivation. *Journal of Personality and Social Psychology*, 37(8), 1352–1363. <https://doi.org/10.1037/0022-3514.37.8.1352>
- Harackiewicz, J. M., Manderlink, G., & Sansone, C. (1984). Rewarding pinball wizardry: Effects of evaluation and cue value on intrinsic interest. *Journal of Personality and Social Psychology*, 47(2), 287–300. <https://doi.org/10.1037/0022-3514.47.2.287>
- Hidi, S. (2016). Revisiting the role of rewards in motivation and learning: Implications of neuroscientific research. *Educational Psychology Review*, 28(1), 61–93. <https://doi.org/10.1007/s10648-015-9307-5>
- Jaeggi, S. M., Buschkuhl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory*, 18(4), 394–412. <https://doi.org/10.1080/09658211003702171>
- Kruglanski, A. W., Fishbach, A., Woolley, K., Bélanger, J. J., Chernikova, M., Molinario, E., & Pierro, A. (2018). A structural model of intrinsic motivation: On the psychology of means-ends fusion. *Psychological Review*, 125(2), 165–182. <https://doi.org/10.1037/rev0000095>
- Laine, T. H., & Lindberg, R. S. (2020). Designing engaging games for education: A systematic literature review on game motivators and design principles. *IEEE Transactions on Learning Technologies*, 13(4), 804–821. <https://doi.org/10.1109/TLT.2020.3018503>
- Lakens, D., & Caldwell, A. (2020). *ANOVA_power*. http://shiny.ies.tue.nl/anova_power/
- Lepper, M. R. (1998). A whole much less than the sum of its parts. *American Psychologist*, 53(6), 675–676. <https://doi.org/10.1037/0003-066X.53.6.675>
- Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology*, 28(1), 129–137. <https://doi.org/10.1037/h0035519>
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. SAGE Publications.
- Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods*, 49(2), 433–442. <https://doi.org/10.3758/s13428-016-0727-z>
- Mekler, E. D., Brühlmann, F., Tuch, A. N., & Opwis, K. (2017). Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. *Computers in Human Behavior*, 71, 525–534. <https://doi.org/10.1016/j.chb.2015.08.048>
- 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 of the United States of America*, 107(49), 20911–20916. <https://doi.org/10.1073/pnas.1013305107>
- Novak, D., Nagle, A., Keller, U., & Riener, R. (2014). Increasing motivation in robot-aided arm rehabilitation with competitive and cooperative gameplay. *Journal of Neuroengineering and Rehabilitation*, 11(1), 64. <https://doi.org/10.1186/1743-0003-11-64>
- R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Reiss, S. (2005). Extrinsic and intrinsic motivation at 30: Unresolved scientific issues. *The Behavior Analyst*, 28(1), 1–14. <https://doi.org/10.1007/BF03392100>
- Rosenfield, D., Folger, R., & Adelman, H. F. (1980). When rewards reflect competence: A qualification of the overjustification effect. *Journal of Personality and Social Psychology*, 39(3), 368–376. <https://doi.org/10.1037/0022-3514.39.3.368>
- Rummel, A., & Feinberg, R. (1988). Cognitive evaluation theory: A meta-analytic review of the literature. *Social Behavior and Personality*, 16(2), 147–164. <https://doi.org/10.2224/sbp.1988.16.2.147>
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Ryan, R. M., & Deci, E. L. (2005). *Intrinsic motivation inventory*. <https://selfdeterminationtheory.org/intrinsic-motivation-inventory/>
- Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of Human-Computer Studies*, 74, 14–31. <https://doi.org/10.1016/j.ijhcs.2014.09.006>
- Shah, J. Y., & Kruglanski, A. W. (2000). The structure and substance of intrinsic motivation. In *Intrinsic and extrinsic motivation* (pp. 105–127). Elsevier. <https://doi.org/10.1016/B978-012619070-0/50027-1>
- Smith, T. W., & Pittman, T. S. (1978). Reward, distraction, and the overjustification effect. *Journal of Personality and Social Psychology*, 36(5), 565–572. <https://doi.org/10.1037/0022-3514.36.5.565>
- Tang, S.-H., & Hall, V. C. (1995). The overjustification effect: A meta-analysis. *Applied Cognitive Psychology*, 9(5), 365–404. <https://doi.org/10.1002/acp.2350090502>
- Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). *Mediation: R package for causal mediation analysis*.
- Unsworth, N., & Spillers, G. J. (2010). Variation in working memory capacity and episodic recall: The contributions of strategic encoding

- and contextual retrieval. *Psychonomic Bulletin & Review*, 17(2), 200–205. <https://doi.org/10.3758/PBR.17.2.200>
- Wiersma, U. J. (1992). The effects of extrinsic rewards in intrinsic motivation: A meta-analysis. *Journal of Occupational and Organizational Psychology*, 65(2), 101–114. <https://doi.org/10.1111/j.2044-8325.1992.tb00488.x>
- Wouters, P., Van Nimwegen, C., Van Oostendorp, H., & Van Der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, 105(2), 249–265. <https://doi.org/10.1037/a0031311>
- Zhang, Y., Fishbach, A., & Kruglanski, A. W. (2007). The dilution model: How additional goals undermine the perceived instrumentality of a shared path. *Journal of Personality and Social Psychology*, 92(3), 389–401. <https://doi.org/10.1037/0022-3514.92.3.389>

Received April 20, 2021

Revision received July 23, 2021

Accepted September 18, 2021 ■