THE NATURE AND EXTENT OF POST-REWARD CROWDING-OUT: THE 'EFFORT-BALANCING' ACCOUNT

INDRANIL GOSWAMI

(igoswami@chicagobooth.edu)

OLEG URMINSKY

(oleg.urminsky@chicagobooth.edu)

*** Please contact the authors for an updated version before citing ***

Authors' Note:

The authors thank Richard Hahn for his valuable statistical insights. The authors would also like to thank Jean-Pierre Dubé, George Wu, Dan Bartels, Abigail Sussman, and seminar participants at the University of Chicago for their helpful comments and feedback on previous versions of this paper. Seed grant from the University of Chicago BIG Ideas is gratefully acknowledged.

ABSTRACT

Although incentives can be a powerful motivator of behavior, research on intrinsic motivation has suggested that rewards can crowd-out task interest, reducing engagement when rewards end. This research has resulted in widespread skepticism among practitioners and academics alike about using incentives in interventions. However, recent field studies examining the long-term effects of temporary incentives have *not* found such effects. We propose a new *Effort-Balancing* account, which suggests that post-reward crowding-out often represents the need for a 'break' after investing effort, rather than a change in interpretation of the task or beliefs about own preferences. As a result, post-reward crowding-out is relatively *momentary* and consistent with longer-term neutral or positive spillover effects of temporary incentives. We test a series of novel predictions, including that momentary crowding-out will be reduced when efforts are more balanced by rewards (e.g., higher reward magnitude or a less effortful activity) or when people are given a break. The implications of momentary crowding-out for long-term behavior and design of incentive programs are discussed.

Keywords: intrinsic motivation, crowding-out, incentives, long-term behavior, decision making

"What rewards do, and what they do with devastating effectiveness, is smother people's enthusiasm for activities they might otherwise enjoy"

-- Alfie Kohn, author of Punished by Rewards

Incentives can be a powerful tool to influence behavior, in settings as diverse as health, education, employment, and marketing promotions (Gneezy, Meier, and Rey-Biel, 2011; Prendergast, 1999; DelVecchio, Henard, and Freling, 2006) where people often make choices involving more immediate or delayed benefits (Loewenstein, Brennan, and Volpp, 2007). Providing immediate temporary incentives can help motivate people to take beneficial action, countering the effects of hyperbolic discounting and present bias (Ainslie, 1975a; Urminsky and Zauberman, 2015) . However, as illustrated by the quote above from prominent educational-policy author Alfie Kohn, the use of incentives has been widely criticized. Specifically, critics argue that economic incentives can undermine, or 'crowd-out' intrinsic motivation and therefore can result in the incentivized behavior becoming less, rather than more prevalent (Kohn, 1999; Pink, 2011; Sandel, 2012), once the temporary incentive ends.

This view stems from a large and influential academic literature on Cognitive Evaluation Theory (Deci and Ryan, 1985) and the Overjustification Hypothesis (Lepper and Greene, 1978), which has studied the effect of temporary incentives (e.g., payments earned during a temporary period for working on a task), and found that task engagement was reduced (e.g., relative to an non-incentivized control group) immediately after the incentive had ended. A highly influential meta-analysis summarized the findings of this literature by warning that "if people use tangible [i.e., monetary] rewards, it is necessary that they be extremely careful… about the intrinsic

motivation and *task persistence* of the people they are rewarding" (Deci et al. 1999, p. 656, emphasis added).

Most of the lab research on post-reward crowding-out with adults has examined the effect of rewards on behavior only immediately after the rewards have been removed. Even though the persistence of post-reward crowding-out has received little empirical scrutiny, widespread skepticism about the long-term consequences of using incentives, generalized from this literature, have played a central role in debates among policy makers about the use of incentives in consequential domains (Lacetera, Macis, and Slonim, 2013). However, recent program-evaluation studies that also followed up on the long-term downstream effects after the incentives were stopped, do not reveal any evidence of post-reward crowding-out, with some even reporting a positive behavioral spillover (e.g., workplace smoking cessation program, Halpern et al., 2015; student incentives for passing exams, Jackson, 2010).

In this paper, we develop a new experimental paradigm to facilitate dynamic measurement of crowding-out. We propose a new "Effort-Balancing" account, which suggests that the post-reward crowding-out effect often represents a need for a 'break' after exerting effort, rather than a change in either task perception or in beliefs about own preferences, as previously proposed. As a result, post-reward crowding-out is relatively *momentary* and is consistent with neutral or even positive overall longer-term spillover effects of temporary incentives. Across five studies, we find a robust momentary crowding-out effect, and test a series of novel predictions, including that momentary crowding-out will be reduced when the efforts are more balanced by rewards (e.g., higher reward magnitude or a lower-effort activity) or when people are given a break.

THEORETICAL DEVELOPMENT

In this paper, we investigate people's behavior *after* a temporary monetary incentive ends. Consider the situation, illustrated in Figure 1, in which people face a series of choices between doing a target task and an alternative option (e.g., a leisure activity). People's choices of how much of the target task to do are based on their intrinsic motivation during the initial pre-incentive period. Some people are then provided with a temporary incentive for doing the target task during the second period, which typically increases effort on that task (but see Gneezy and Rustichini, 2000b for some exceptions). If, after the incentive ends, effort on the target task falls below what it would have been without the incentive (e.g., the control group), post-reward crowding-out has occurred. The magnitude and duration of post-reward crowding-out, relative to any increased effort during the incentive period, determines the net effect of the incentive.

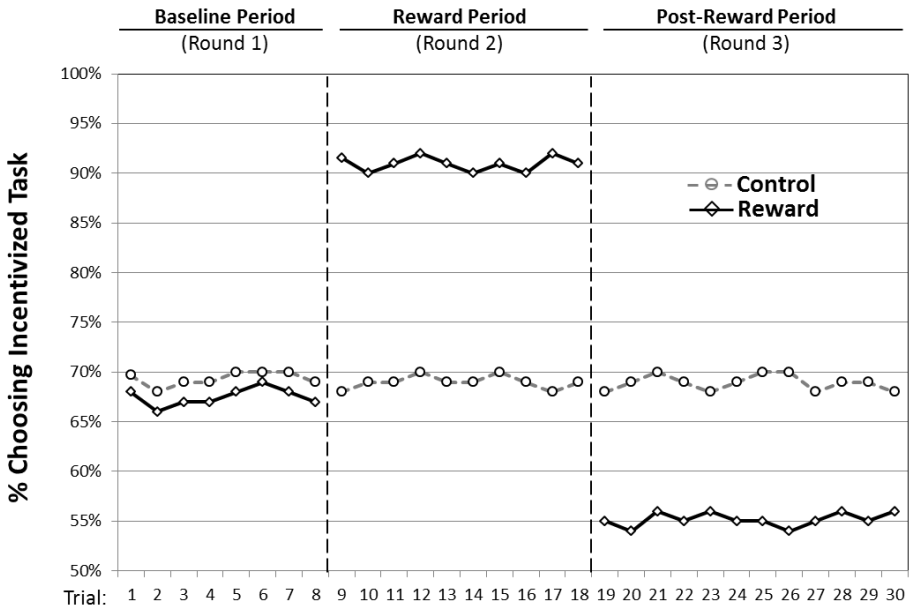

Figure 1: Schematic illustration of persistent post-reward crowding-out. After the incentive ends, theories of 'crowding-out' predict that effort will fall below the pre-incentive level. The scope and duration of crowding-out in Round 3 relative to the increase in Round 2 determines the net effect of incentive.

The post-reward crowding-out illustrated in Figure 1 is implied by existing theories of post-reward crowding-out. Next, we review the competing theories and their predictions for the nature and extent of post-incentive motivation in more detail. Then, we specify the distinct predictions made by our Effort-Balancing account.

EXISTING THEORIES OF POST-REWARD CROWDING-OUT

Undermining Autonomy. Cognitive Evaluation Theory argues that people engage in tasks which fulfill innate needs of competence, exploration, and autonomy (White, 1959; Deci, 1971). An external reward to do a task, by exerting external control over people's behavior, is posited to undermine satisfaction from autonomy and exploration (Deci and Ryan, 1985; Ryan and Deci, 2000) and therefore to make the task incapable of fulfilling these intrinsic needs. As a result, people often will do the task for the incentive, but engagement in the task will fall below the initial baseline once incentives end (unless the incentive communicates competence, which can countervail the negative effects). This account predicts that because people's interpretations of the task are changed by incentives, the post-reward crowding-out effect will be persistent (Deci et al. 1999), and will be more evident when the task is more intrinsically rewarding or the incentive is larger.

Overjustification. An alternative self-perception account (Bem 1967) focuses on the effect of incentives on inferences about one's own preferences, rather than views of the task. When a task is incentivized, people may infer that they were doing the task only because of the external rewards (Kruglanski et al., 1972; Lepper et al., 1973; Lepper and Greene, 1978), discount their own intrinsic motives (Nisbett and Valins, 1971; Kelley, 1973) and infer that they

do not like doing the task. For example, providing a sufficiently large temporary price discount could reduce post-promotion sales of a product, because people assume that they only purchased because of the discount and consequently infer a lower level of product liking (Dodson, Tybout, and Sternthal, 1978). Because the post-reward crowding-out is due to a change in people's beliefs about their own preferences, over-justification predicts that the reduction will persist until another influence causes those preferences to change, and the reduction will be stronger for higher-magnitude incentives. However, unlike Cognitive Evaluation Theory, Overjustification does not make a clear prediction about post-reward behavior if people are temporarily incentivized for a more versus less interesting task.

Task and Social Inferences. People may also infer unknown aspects of the task from an incentive by considering the incentive provider's motives. Since incentives are often introduced when tasks are believed to be unattractive or difficult, people may attribute such beliefs to the individual who decided to offer the incentive (Benabou and Tirole, 2003). Furthermore, small incentives could also lead to an inference that the incentivized behavior is unimportant (Gneezy and Rustichini, 2000a). Consumers are known to make such inferences about missing information in the marketplace (Huber and McCann, 1982), such as price-quality associations (Levin, Johnson, and Faraone, 1984), and certain promotions can therefore have a negative impact on sales (Simonson et al., 1994).

As a result, people may put less effort into the task when the incentive ends, because of negative inferences about task importance, enjoyment or difficulty, for as long as those inferences persist. Notably, such inferences can also reduce the desired behavior during the period when incentives are available (Frey and Oberholzer-Gee, 1997; Fehr and Falk, 2002; see Frey and Jegen, 2001 for a review), resulting in *during-reward* crowding-out (as opposed to

post-reward crowding out). These accounts suggest that prior experience with the task would reduce crowding-out due to inference making (both during-reward and post-reward), and both low and higher incentives could result in crowding-out.

Expectations. Unexpectedly ending an incentive can be seen as a violation of norms or expectations, resulting in disappointment (Fehr and Falk, 2002). Relatedly, the incentive may create a reference point, such that then doing the task without an incentive could be viewed as a loss (Kalyanaram and Winer, 1995). In these accounts, when people learn that their expectation of continuing incentives will not be fulfilled, engagement in the task can fall (Esteves-Sorensen, Macera, and Broce, 2013). Since people cope with disappointment and adjust their reference points over time (Gilbert, Pinel, Wilson, Blumberg, and Wheatley, 1998; Schkade and Kahneman, 1998), these accounts predict that post-reward crowding-out will be temporary, and will be eliminated when people have advance knowledge of the temporary nature of the incentive.

Temporal Shift in Effort. When decision makers know that the incentive will be temporary, they can move up the effort they would have invested during the post-incentive period, instead doing more of the task while the incentive is available. In such cases, there might be no change in the overall net effect of incentives, but an increase during the incentive period and a decrease afterwards, which could appear to be post-reward crowding-out. For example, employees might move up task completion in time to help them earn a bonus (Oyer, 1998) or people might stockpile non-perishable goods while a purchase incentive is available (Mela, Jedidi, and Bowman, 1998). Crowding-out due to effort shifting should occur primarily when people know about the availability and timing of post-reward tasks. The post-reward

consequences of temporal shifting should be larger and persist for longer when higher incentives are provided.

In this paper, we will study the effect of incentives on *post-reward* motivation and behavior. Incentives can sometimes also reduce engagement in the incentivized task while available (i.e., *during-reward*). This can occur when incentives convey missing information about the task or decision environment, or if the benefits from engaging in the task largely accrue from signaling motives (Bénabou and Tirole, 2005; Ariely, Bracha, and Meier, 2009; Dube, Xueming, and Fang, 2015). In particular, when small incentives are used, the low value of the incentive does not counter the detrimental effect on motivation of introducing external motives (Gneezy and Rustichini, 2000b). While such failed incentives could also affect post-reward behavior, our focus in this paper is on how *successful* incentives (i.e., those that increase task effort while available) affect behavior when they end. Accordingly, we do not study pro-social tasks (which have high signaling value) or situations with high uncertainty, requiring additional inferences.

Most of the accounts, as described above, predict persistent post-reward crowding-out from introducing temporary incentives that are successful during-reward. However, as mentioned earlier, recent field studies tracking post-program behavior of temporary incentive policies have failed to find such long-term negative effects. Next, we propose an alternative account of post-reward crowding-out effects, which can reconcile the disparate findings in the literature and makes several novel testable predictions.

THE ‘EFFORT-BALANCING’ ACCOUNT: CROWDING-OUT AS TAKING A BREAK

Consider activities like going to the gym, going to the doctor for a regular health checkup, learning a new skill or practicing an existing skill. Even though these tasks (sometimes called “virtuous”, in a far-sighted rather than moral sense e.g., Hoch and Loewenstein, 1991) require effort, people often feel they *should* be doing such activities and regularly engage in these tasks without any external reward, reflecting an intrinsic motivation. When people have the goal to engage in such effortful activities, they may at the same time also have a goal to seek leisure. People often strive to find the right balance between competing goals (Urminsky and Kivetz, 2011), including between effort and leisure (Kool and Botvinick, 2014), and between effort and expected reward (Kivetz, 2003).

When temporary incentives are introduced for a virtuous effortful task, the reasons for pursuing the incentivized goal are bolstered, as the more immediate incentive of earning rewards is added to the intrinsic motivation of pursuing a desirable or beneficial activity. When goals compete for resources, people are more motivated by the goals that are more salient (Shah and Kruglanski, 2002) and more proximal (Brown, 1948; Kivetz, Urminsky, and Zheng, 2006; Urminsky and Goswami, 2015), which would favor the currently incentivized task over leisure.

Thus, while pursuing the incentivized task, leisure goals are likely to be deferred to avoid interruption (Jhang and Lynch, 2015) until a time deemed more appropriate (Shu and Gneezy, 2010). As people forego leisure in favor of the compensated virtuous work, the leisure goal is likely to nevertheless persist in the background. Thus, when the incentive ends, motivation will shift from the completed incentive-earning goal, and thereby from the virtuous effortful task, to the neglected competing goal (Fishbach and Dhar, 2005) of pursuing leisure. As a result,

immediately after the incentive ends, the goal associated with the effortful task will be weakest (Kruglanski et al., 2002; Khan and Dhar, 2006), and people will be motivated to engage in the leisure task.

Furthermore, the disengagement from the incentivized task in favor of the leisure task is likely to be *temporary*. Even minimal satisfaction of an immediate goal can restore motivation for longer-term goals (Urminsky and Kivetz 2011). More generally, people adapt to changes in the environment (Gilbert et al., 1998; Schkade and Kahneman, 1998) and spontaneous recovery of previously reinforced behavior sometimes happen merely from the passage of time (Brooks and Bouton, 1993). Thus, after engaging in a leisure task for a while, the preferred balance between effort and leisure is likely to be restored (Kool and Botvinick, 2014), resulting in people's behavior returning back to their baseline level of preference.

Based on this "Effort-Balancing" account, we propose that incentivizing an intrinsically motivating but effortful task is likely to result in "momentary" post-reward crowding-out, occurring immediately after the incentives end, but only persisting temporarily.

> **H1a:** Post-reward crowding-out happens for a short duration immediately after the incentive ends, with post-reward effort subsequently returning to or exceeding the baseline level.

In our Effort-Balancing account, post-reward momentary crowding-out is best understood as taking a 'break' from the incentivized work, rather than a persistent change in beliefs, either about own preferences (Lepper and Greene, 1978) or about the task (Deci and Ryan, 1985). Therefore, unlike prior theories, we suggest that taking a break will reduce crowding-out.

H1b: Providing an explicit non-effortful break from the target task, immediately after the incentive ends, will reduce post-reward crowding-out.

Our Effort-Balancing account, in characterizing post-reward crowding-out as a break from effortful work, suggests an important moderator. Effort can be balanced by either leisure (e.g., taking a break) or by rewards. Choices between virtues and relative vices are affected by how justified people feel in choosing one over the other (Kivetz and Zheng, 2006; Sela, Berger, and Liu, 2009). Although prior efforts provide a justification for choosing leisure, high rewards can balance out that effort, reducing the justification for leisure. Therefore the more rewarding the incentive, relative to effort exerted, the less of a justification people will feel for engaging in the leisure activity, reducing the motivation to take a break. In such situations, we predict that momentary crowding-out would be reduced.

Furthermore, the new account makes an important prediction about when temporary incentives might even result in a positive longer-term post-reward spillover effect, complementing the positive during-reward effects of incentives. We propose that the *direction* of the overall post-reward effect of temporary incentives will be moderated by the relative magnitude of the rewards. Research on evaluating conditioning (Razran, 1954) has suggested that when rewarding outcomes are associated with a task, people like the task more, even when those rewarding outcomes are no longer present, because of the positive valence transfer (De Houwer et al., 2001). Thus, when the incentive is highly rewarding relative to the effort invested, not only will momentary post-reward crowding-out be reduced or eliminated, but the incentive could result in greater liking of the incentivized task and post-reward crowding-in (i.e., positive

spillover). In this case, engagement would persist at a higher level than the baseline, at least for some time after the rewards are withdrawn.

> **H2:** Higher perceived incentive magnitude will reduce post-reward momentary crowding-out, and can even yield post-reward crowding-in, increasing overall post-reward task effort.

It is important to note that this hypothesis is the opposite of what would be predicted by prior theories of post-reward crowding-out. Specifically, prior accounts suggest that since larger incentives exert a stronger influence over behavior, larger incentives are more likely to undermine feelings of autonomy and provide a stronger external attribution for effort, yielding a larger crowding-out effect. Similarly, larger temporary incentives would yield higher reference-point based expectations, and stronger post-reward crowding-out. Thus, varying incentive magnitude serves as a strong test between our account and prior accounts.

Finally, the proposed account also makes unique predictions when a less effortful activity is incentivized. Since adding an incentive to a more leisure-like activity makes the typically low-effort experience even more rewarding, people are less likely to feel that the effort balance needs to be corrected. Therefore, we predict that incentivizing leisure tasks will be less likely to result in post-reward crowding-out. In contrast, the existing theories of crowding-out would predict that providing incentives for leisure tasks would be more detrimental, yielding more crowding-out than for a work task (Calder and Staw, 1975), since the less effortful, more leisure-like activities are likely to be interesting and intrinsically motivating to people.

H3: Incentivizing a leisure activity will result in less post-reward crowding-out than equivalent incentives for a more-effortful task.

We describe five experiments and a combined meta-analysis to test our proposed account. In Study 1, we introduce the experimental setup that we will use in the subsequent studies, and provide evidence that crowding-out is relatively momentary, as predicted by the Effort-Balancing account. Merging equivalent conditions from all studies conducted into a single meta-analysis, we find that the momentary crowding-out effect is highly robust. In Study 2, we find that giving people an explicit `break' from the incentivized task after the reward ends can preclude post-reward crowding-out. We then test the effects of reward magnitude (Study 3) and framed task importance (Study 4) on the momentary crowding-out effect. Finally, we vary the type of target task, and find that, contrary to the prediction of the current theories but consistent with our proposed Effort-Balancing account, the momentary crowding-out observed when incentivizing a work task is eliminated when the leisure task is incentivized instead (Study 5). Across the studies, we contrast our findings with prior theories of crowding-out, and rule out multiple alternative explanations for our findings, including depletion, fatigue, variety-seeking, social-desirability, and signaling motives.

STUDY 1: MOMENTARY CROWDING-OUT

Method

We designed a flexible experimental paradigm which enables us to investigate the dynamic effects of incentives, including the initial effects on post-reward behavior, the

persistence of any post-reward crowding-out, and the longer-term post-reward effort level (e.g., spillover).

Participants made repeated choices (illustrated in Figure 2) between doing an effortful math problem (work task) or watching a brief video (leisure task) in the next trial, divided into three multiple-trial rounds: pre-incentive, during-incentive, and post-incentive[1]. The math task, adapted from Experiment 1 of Mazar, Amir, and Ariely (2008), involved finding the two numbers which add up to 10 in a grid of 12 numbers, and had to be completed within a 30 second time limit. Participants were told that every math task had only one correct answer. The video clips were also 30 seconds long in duration.

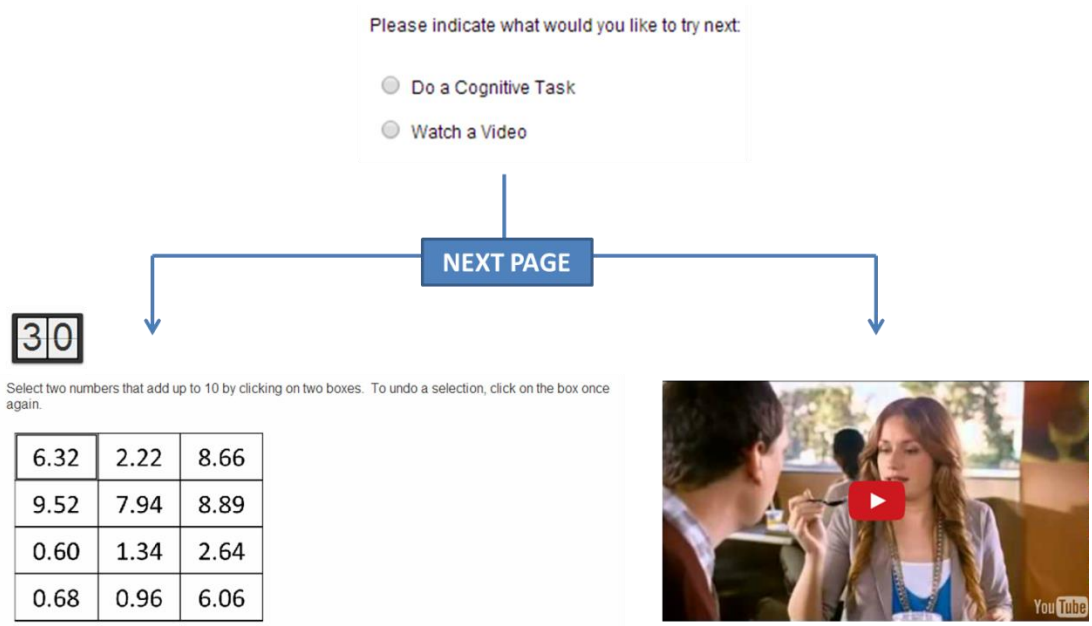

Figure 2: The figure shows a typical choice participants made in each trial. The outcome of the choice (i.e., which task or which video) was revealed on the next page. The math task involved finding two numbers that add up to 10 and each grid had only one solution. In the actual experiment, each participant made 30 such sequential choices.

[1] Detailed stimuli are available in the Appendix A of the Online Supplemental Materials.

All participants were given the following instructions:

"In this survey you will be a given a series of choices between doing cognitive tasks and watching videos of interesting television advertisements collected from across the world.

The cognitive task will train your mental reasoning skills, and we will use your results to calibrate and standardize a training test. You can do as many of them as you want, or can just enjoy the videos."

A pretest (N=47) confirmed that a random sample of participants from the same population judged the math task more work-like compared to the video ($M_{math} = 6.87\ vs. M_{video} = 2.49; t(46) = 9.7, p < .001$), on 9-point scales, but considered the video more leisure-like compared to the math ($M_{math} = 3.89\ vs. M_{video} = 6.78; t(46) = 6.2, p < .001$). The math task was also judged as more effortful ($M_{math} = 5.59\ vs. M_{video} = 1.95; t(46) = 9.0, p < .001$) and less entertaining compared to the video ($M_{math} = 4.64\ vs. M_{video} = 6.43; t(46) = 4.2, p < .001$). Consequently, participants in the pre-test felt that the math task had more long-term benefits whereas the videos had higher immediate benefits ($M_{math} = 5.38\ vs. M_{video} = 3.97; t(46) = 3.6, p < .001$).

The pre-test also confirmed that choosing the math task over the video task was seen as relatively more virtuous (on a 1=Vice to 9=Virtue scale) than choosing the video over the math task ($M_{choose\ math} = 6.00\ vs. M_{choose\ video} = 4.43; t(46) = 3.7, p < .001$). Likewise, more justification was needed for choosing to watch the video task over doing the math tasks, than for the opposite choice ($M_{choose\ math} = 2.97\ vs. M_{choose\ video} = 4.81; t(46) = 3.9, p < .001$).

Thus, the choice between the doing a math task and watching a video represented a self-control dilemma. Participants chose between a relatively virtuous option that entailed exerting effort but could train their mental reasoning skills, and a less virtuous leisure option that involved

watching interesting videos. The specific task to be performed in the trial (i.e., the exact video or the numbers in the math task) was revealed only after the choice was made.

Finally, an important pre-requisite of any framework claiming to investigate the effect of external incentives on intrinsic motivation is to use tasks that are high on self-rated interest and liking. A commonly accepted criteria (Deci, Koestner, and Ryan, 1999) is for the target-task to have a rating of at least the mid-point on a scale measuring how "interesting and enjoyable" the task is (1=Low, 9=High). The pre-test confirmed that both the math task ($M_{math} = 5.85$) and the video task ($M_{video} = 6.27$) fulfilled this requirement.

Our experimental paradigm was specifically designed to capture dynamic changes in behavior over time, even after a temporary disengagement. It is important to note that the "quitting paradigm" commonly used to measure task persistence (Deci, 1971; Heyman and Ariely, 2004) would not have provided the necessary moment-by-moment tracking of engagement. In the quitting paradigm, participants engage in an unsolvable task and persistence is measured by time spent before quitting the task. However, since there is no opportunity for re-engagement, initial disengagement could be mistaken for a persistent propensity to disengage.

However, quitting could potentially be an issue in our setup as well, which we address in several ways. Participants could "quit" within the study, by repeatedly deciding to only watch the videos (the leisure activity) for the remaining duration of the study. However, participants could also quit by ending the study and not completing the remaining tasks. We tracked all drop-outs, and included participants who dropped-out of the study after completing the pre-reward baseline period (around 5% of the total participants) in all of our analyses, coding their participation as

zero for the target task. This ensured that we included anyone whose behavior could have been impacted by the incentives, whether they finished the study or not[2].

Online participants (N=77) were randomly assigned to one of two conditions – the control condition or the incentive condition. After an initial sample trial (a pre-solved math-task and a sample video), we confirmed that the participants had no technical problems watching the video, and they began the first round. In both conditions, participants completed eight trials in the initial round, first making their choice and then engaging in the chosen task in each trial.

Next, participants in the incentive condition were informed that they would earn 5 cents for every correct answer (i.e., a performance-contingent incentive), but only during the next round, and that no such rewards would be available in any subsequent rounds. Participants in the control group proceeded to the second round without any additional information. After completing Round 2, which consisted of 10 trials, participants in the incentive condition were given feedback about their performance on the math task and told how much incentive they had earned. Participants in both conditions then moved on to the third round, which consisted of 12 trials. After completing demographic questions, participants in both groups were paid a nominal participation fee (plus any incentives earned by the reward group). Both the math tasks and the videos took the same amount of time (30 seconds) and, as a result, the total time to complete the study was unrelated to participant's choices between tasks.

The experimental paradigm allows us to distinguish between the effects of incentives on effort (i.e., choices of the math task), accuracy (probability of correctly solving the math task), and net outcome (total number of correct answers). We will primarily focus on effort (e.g.,

[2] We examine robustness by excluding these dropouts in the meta-analysis and the results do not change, as reported in Appendix B of Online Supplemental Materials.

choosing to do the target task) as the key variable of interest, consistent with the approach used in the intrinsic motivation literature. Accuracy on the math task did not significantly differ across conditions in any of the studies, and therefore differences in net outcomes (number of correctly solved math problems) were determined by effort.

Results

In the control condition, participants chose the math task 67% of the time during the baseline period. The average effort (i.e., choices of the math task) remained at around the same level during the following experimental periods in the control condition, as shown in Figure 3. The willingness of the participants to engage in the math task without any additional incentive is important, confirming that the task was intrinsically motivating (Deci, Koestner, and Ryan, 1999). Tasks that have little or no baseline engagement cannot show substantial crowding-out, by definition, because there is no intrinsic motivation to be affected by incentives.

In the incentive condition, the proportion of math task choices increased to 88% when the incentive was available – a significant increase compared 61% in the control condition during the same period, controlling for the baseline proportion of math task choices in Round 1[3](t=5.4, p<.001). After the incentive ended at the conclusion of Round 2, the average percentage of math task choices in the reward condition during the post-reward Round 3 was very similar to that in the control condition during the same period (66% vs. 65%; $t < 1$). Thus, there was no overall post-reward crowding-out of effort due to the temporary incentive. These results are consistent with recent field studies that have also not found long-term negative effects of temporary incentives (Chetty, Saez, and Sándor, 2014; John et al., 2011; Volpp et al., 2006).

[3] In all studies the differences in individual-level proportions of target task choices are reported along with linear regression t-tests that control for the individual baseline levels in Round 1. We also report results from parametric hierarchical regression models.

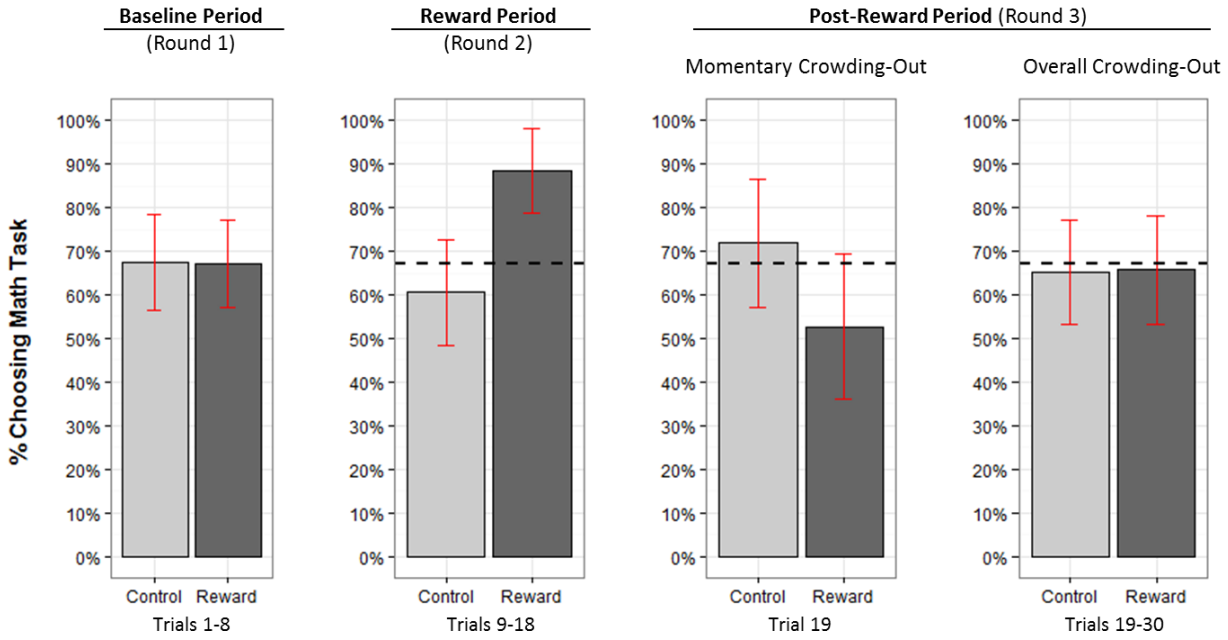

Figure 3: Average proportion of math task choices in the control and reward group in each phase of the experiment. The first graph shows the baseline period. The second graph shows the increase in the reward group due to incentives. The third graph shows the decrease in effort for the reward group in the first post-reward trial (momentary crowding-out). The fourth graph shows no overall post-reward effect of incentives. The dotted lines represent the average effort level across the conditions in the baseline (i.e., pre-incentive) period, and the vertical lines are 95% CIs.

However, because our experimental paradigm allows us to identify variation in engagement over time, we can distinguish behavior immediately after the incentive ended from overall post-reward behavior. In the very first trial in Round 3, the average percentage of math task choices dropped to 53%, 19 percentage-points lower than in the control condition (72%), a significant difference controlling for the baseline proportion of math task choices in Round 1 ($t = 2.0, p = .05$). This immediate drop in people's choices of the math task after the incentived Round 2 concluded is consistent with prior lab studies which had observed reductions in the incentivized behavior immediately after rewards ended (Deci, 1972; Calder and Staw, 1975; Pittman, Davey, Alafat, Wetherill, and Kramer, 1980; Ryan, Mims, and Koestner, 1983).

Thus our findings reveal a *momentary* post-reward crowding-out effect. Because crowding-out was only momentary, the negative post-reward effect of incentives was small in scope relative to the positive during-reward effect of incentives. Overall, the use of incentives in the reward condition yielded a significant positive net effect in the study, with more choices of the math task in the incentive condition than in the control condition during the post-baseline periods, controlling for Round 1 (76% vs. 63%; t = 2.9, p=.005).

Mixed Model for Momentary Crowding-out Effects

Using linear regression t-tests we tested the effect of incentives on choices in the first post-reward trial only. However, our hypotheses about the persistence of crowding-out post-reward require a more flexible test that allows us to quantify the combined effect of momentary crowding-out, which may or may not extend for multiple trials. Therefore, to more precisely quantify and test the magnitude of momentary crowding-out, we fit a hierarchical non-linear mixed model (Raudenbush and Bryk, 2002), with the per-trial observations nested under individuals.

In our model, we capture total momentary post-reward crowding-out using a functional form assumption about how effort returns to baseline over time in the post-reward period. Assuming a non-linear return of effort (i.e., likelihood of choosing the math task) to baseline over time (the number of periods t since the incentive ended), we parameterize momentary post-reward crowding-out (MCO) as:

$$MCO_t = \frac{1}{t} \tag{1}$$

Using this parameterization, the probability of individual i choosing to do the math task in post-reward (Round 3) during trial t can be written as:

$$P(Y_{ti} = 1) = \phi\ (\beta_{0i} + \ \beta_{Mi}\ MCO_t) \tag{2}$$

In our model, we set ϕ to the logit link function. In the hierarchical regression, the parameters in Equation (2) are a function of time-invariant individual-level covariates, to account for the repeated observations per person.

$$\beta_{0i} = \beta_{00} + \ \beta_{01}\ C_i + \ \beta_{02}\ X_i + u_{0i} \tag{3}$$

The person-specific baseline likelihood parameter β_{0i} is a function of the condition that individual i has been randomly assigned to (C_i), the total number of math task choices by individual i in the pre-incentive Round 1 (X_i), as well as the population baseline β_{00} and time-invariant person-specific error term u_{0i}.

$$\beta_{Mi} = \beta_{10} + \ \beta_{11}\ C_i + \ u_{1i} \tag{4}$$

The person-specific momentary crowding-out effect, β_{Mi}, is estimated as a function of experimental condition C_i, as well as the baseline momentary crowding-out β_{10} and the individual-specific error term u_{1i}. The random effects for the intercept and the slope for every individual i, u_{0i}, u_{i1}, are assumed to be bi-variate normal with zero-mean, variances τ_{00}, τ_{11} and common co-variance τ_{01}. This error structure accounts for the potentially correlated repeated-measures for each individual. Combining equations, (2), (3), and (4) yields an "intercepts and slopes-as-outcomes" model (Raudenbush and Bryk, 2002). The proportion of math tasks chosen in each trial t of Round 3 is:

$$P(Y_{ti} = 1) = \phi\ (\beta_0 + \ \beta_1 MCO_t + \ \beta_2 C_i + \beta_3\ X_i + \ \beta_{MCO}\ C_i * MCO_t) \qquad (5)$$

Thus, β_{MCO} tests for a difference in the extent of momentary crowding-out in the experimental condition (C_i=1), compared to the corresponding time periods in the control condition (C_i=0). A significant and negative β_{MCO} generally indicates momentary crowding-out after the incentive ended, compared to the corresponding trials in the control condition, controlling for individual differences in baseline effort X_i. However, an estimate of β_{MCO} that is not statistically distinguishable from zero represents a consistent level of effort throughout Round 3, with two very different potential interpretations. A non-significant β_{MCO} could indicate that no post-reward crowding-out has occurred or it could represent a consistent overall post-reward spillover (positive or negative) in Round 3. Furthermore, a significant negative β_{MCO} will not necessarily represent initial crowding-out if there is overall crowding-in (i.e., high levels of math task choices during Round 3). Thus, it will be important to also estimate the overall effects of incentives on choices in Round 3, in addition to momentary crowding-out, and interpret these parameters jointly. Next, we describe the tests we use to estimate the overall effects of incentives.

Difference-in-Difference Model for Overall Effects

In addition to the MCO model, we use a hierarchical non-linear difference-in-difference model to estimate differences in the probability of choosing the math task for a given period or periods within the experiment (R_t=1), between two experimental conditions (C_i), controlling for behavior in the earlier period (R_t=0). This general model can be written as follows:

$$P(Y_{ti} = 1) = \phi\ (\beta_0 + \ \beta_1 * \ R_t + \ \beta_2 * C_i + \beta_3 * R_t * C_i\ ) \tag{6}$$

The interpretation of the key coefficient β_3 depends on how the periods (R_t) and conditions (C_i) are coded. Typically, we will compare an incentive condition (C_i =1) to the control condition (C_i =0). To estimate the effect of incentives on during-reward performance in Round 2, β_{REWARD}, we compare during-incentive Round 2 (R_t=1) to baseline Round 1 (R_t=0) and exclude Round 3 data. To estimate the overall crowding-out effect, β_{CO}, we compare post-reward Round 3 (R_t=1) to baseline Round 1 (R_t=0), and exclude Round 2. Similarly, the immediate crowding-out effect in the first post-reward trial, $\beta_{IMMEDIATE}$, is estimated by comparing the first trial in Round 3 (R_t=1) to all of baseline Round 1 (R_t=0), excluding the other trials. Lastly, the net effect of incentives (including during-reward and post-reward effects), β_{NET}, is estimated by comparing the combined during-reward Round 2 and post-reward Round 3 trials (R_t=1) to baseline Round 1 (R_t=0).

Regression Results for Study1

Using the hierarchical regression specification, we find a significant coefficient for momentary crowding-out after rewards had ended ($\beta_{MCO} = \ -2.6, z = -2.6, p < .01$). The reduction in engagement immediately after the incentives were stopped was also significant ($\beta_{IMMEDIATE} = \ -1.3, z = \ -2.0, p = .04$). Moreover, the regression also confirmed that there was no overall crowding-out for the math task from the temporary incentive ($\beta_{CO} = -0.003, z = -0.01, p > .250$). Thus, choices of the math task returned to the baseline level

after the initial post-reward momentary decrease. This suggests that post-reward crowding-out is momentary, and happens temporarily after the incentives have ended.

Overall, since the incentives boosted math task choices significantly over the baseline level during Round 2 ($\beta_{REWARD} = +4.6, z = +4.2, p < .001$), the net effect of incentives was to significantly increase effort on the math task over the course of the study ($\beta_{NET} = +1.2, z = +2.7, p = .006$). Therefore, despite significant momentary crowding-out, providing the incentive increased the number of math tasks that people did in the experiment.

META-ANALYSIS OF THE MOMENTARY CROWDING-OUT EFFECT

We used the same pair of control and reward conditions from Study 1 in all of our studies, including some additional studies conducted for a separate project, and not included in this paper. We conducted a meta-analysis of the basic momentary crowding-out effect using these two focal conditions across all 10 studies and 1225 participants (551 in control, and 674 in the reward group). The large sample size in this internal meta-analysis allows us to not only estimate the effects more precisely but also to perform several robustness checks.

The average percentage of participants choosing the math task in the incentive (solid line) vs. control (dashed line) conditions is shown in Figure 4. The results indicate significant momentary crowding-out in the incentive condition immediately after the rewards have ended, in the first trial of Round 3, relative to the control condition ($\beta_{IMMEDIATE} = -0.5, z = -3.0, p = .002$). Consequently, we find highly significant momentary crowding-out after incentives have ended ($\beta_{MCO} = -2.5, z = -7.5, p < .0001$). This result is robust to different parameterizations

of the MCO variable (e.g., $MCO = t^{-n}$, with different values of $n > 0$). The results are also robust to excluding participants who dropped out before completing the full study[4].

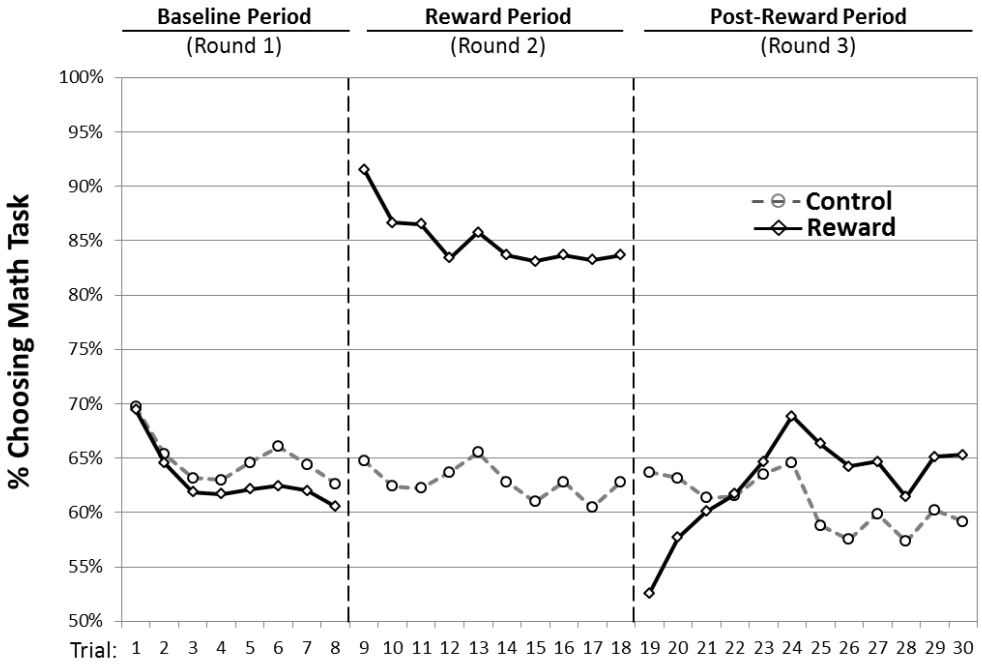

Figure 4: Raw percentage of participants choosing the math task in the various experimental rounds in the combined meta-analysis. There is an immediate but momentary dip in math task choices after incentives were stopped for the reward group, but the choices increase and then settle at a level higher than baseline.

In addition, we modeled post-reward choices of the math task using a non-parametric spline model, and bootstrapped tests[5]. The results suggest that in the first task after the incentive ends, people in the incentive condition have between a 6% and 16% (95% confidence interval) lower probability of choosing the math task, compared to the same trial for the control group. There continues to be a significantly lower level of engagement in the second post-reward choice as well [-2%, -10%]. From the sixth post-reward choice onward, the probability of choosing the

[4] See Appendix B of Online Supplemental Materials for details.
[5] See Appendix D of Online Supplemental Materials for details.

math task is actually significantly higher in the reward group compared to the control group, and this 'crowding-in' continued until the end of the experiment.

Consistent with the findings of this per-trial analysis, we find a significant overall *increase* in choices of the math task post-reward ($\beta_{CO} = +0.4, z = +2.5, p = .02$). This suggests that overall there was a modest crowding-in or positive spillover effect. Furthermore, given the momentary nature of post-reward crowding-out and the otherwise positive effects of incentives, the combined data also show a strong positive net effect of incentives on the number of math tasks chosen ($\beta_{NET} = +1.2, z = +10.0, p <.0001$).

Effort is an important behavioral outcome variable that represents the person's motivation level. Our flexible experimental paradigm allowed us to also examine the effect of temporary incentives on accuracy (probability of correctly answering the math task after deciding to attempt it) and net outcome (total number of correct answers). The incentive could have resulted in people choosing the math task without being able to answer correctly, resulting in a decrease in accuracy compared to the control condition. However, we did not observe any such effects and temporary incentives did not affect accuracy at all. Therefore, the same conclusions hold for effort and net outcome– a significant positive effect of incentives on the total number of math problems solved correctly ($\beta_{NET} = +0.4, z = +4.9, p < .0001$)[6].

Discussion

The findings of Study 1 and the combined data analysis lend support to our proposed Effort-Balancing account, which characterizes post-reward crowding-out as people taking a break after exerting effort (Hypothesis H1a). Our account differs from prior theories, which were

[6] See appendix E of the Online Supplemental Materials.

based on a change in either task interpretation or in beliefs about own preferences, and which implied that people would lose interest in the task because of the incentive, resulting in persistent crowding-out. In fact, we found no differences in self-reported task interest at the end of the experiment (1- extremely uninteresting and unenjoyable to 9 - extremely interesting and enjoyable) between the reward and control groups ($M_{reward} = 6.3\ vs\ M_{control} = 6.1;\ t(71) < 1$) in Study 1, suggesting that the incentive did not reduce self-reported intrinsic motivation.

Our experimental paradigm was also designed to prevent confounds between our measure of crowding-out and several other alternative behaviors. Participants were not told in advance about the upcoming rewards, to avoid affecting the baseline measures of engagement in Round 1. Participants also did not need to make inferences about the task characteristics from the incentives, because they had already experienced the task first-hand for multiple trials during the baseline period, before the incentives were introduced.

At the beginning of the reward-round, the treatment group was also told that no rewards would be available in any subsequent rounds. This was done to prevent post-reward disappointment or reference-point formation, which could reduce motivation if the participants only discovered at the end of Round 2 that the rewards had ended. Informing participants that there would be no further incentives also precluded any post-reward strategic slacking in the hope that rewards could return.

Further, in order to avoid temporal substitution of effort between the during-reward round and the post-reward round, participants in the experiment did not know how many trials would occur in any round, and each round used a fixed but different number of trials. As a result, the reported findings cannot be explained by unfulfilled expectations or disappointment, inference about the target-task, temporal shift in effort, or strategic motives.

Lastly, the pattern of results depicted in Figure 4 is also inconsistent with an 'extinction decay' curve as predicted by the operant-reinforcement literature (Skinner, 1953). Specifically, the operant-reinforcement literature would predict that after the reinforcements (i.e., incentives) are stopped, the effort will gradually decline and return to baseline. However, out Effort-Balancing account predicts the need for a 'break' when the rewards do not seem sufficient to balance the high effort expended. As a result, we would predict a temporary disengagement from the target-task, giving rise to a very different pattern of behavior from that predicted by operant-conditioning theories.

Relating the Results to Past Findings

The vast prior literature on post-reward effects in psychology has investigated behavior immediately after the rewards were stopped. In the widely-used one-session paradigm (Deci, 1972; Calder and Staw, 1975; Pittman, Davey, Alafat, Wetherill, and Kramer, 1980; Ryan, Mims, and Koestner, 1983; Houlfort, Koestner, Joussemet, Nantel-Vivier, and Lekes, 2002) to study the effect of rewards on motivation, subjects in both the control and the reward group were given a presumably intrinsically motivating but effortful task (alternative options included reading recent issues of various magazines) and only the reward group was paid for it. Immediately after this experimental session was over, subjects in both groups were unobtrusively observed during a "free-choice" period. Subjects in the reward group typically spent less time on the target task than subjects in the control group during the same time.

This design has been used repeatedly in the literature and results of the free-choice period behavior during the one post-incentive session have been interpreted as suggesting negative post-

reward effects of temporary incentives on intrinsic motivation. Since participants in the immediate post-incentive free-choice session presumably made one decision whether to engage in the target-task or to engage in a leisure activity instead, this would correspond to the immediate (or first few) post-reward trial(s) in our paradigm. Therefore, the momentary post-reward crowding-out effect in our studies is consistent with the crowding-out effect reported in the prior lab studies with adult participants, which did not capture the dynamic or delayed effects of the incentive[7].

The implication from extant literature, that temporary incentives would have a persistent crowding-out effect, is broadly inconsistent with recent field studies examining the net effects of various behavioral interventions (e.g., providing access to counseling services, commitment contracts, refundable bonds, social incentives) including monetary incentives, on long-term behavior, none of which reported long-term negative effects after the incentive ended (Chetty et al., 2014; Royer, Stehr, and Sydnor, 2012; Volpp et al., 2008). In fact, some studies have reported a modest positive-spillover effect (Charness and Gneezy, 2009; Jackson, 2010; Volpp et al., 2009), and others have reported return-to-baseline effects (John et al., 2011; Lacetera, Macis, and Slonim, 2011; Volpp et al., 2006). These studies were not specifically designed to examine post-reward crowding-out effects, examining tasks that are quite infrequent (e.g., blood donation, submitting referee reports) and measuring total behavior some time (e.g., days or weeks) after stopping the incentives. However, our studies, which do not find overall crowding-out of temporary rewards, are also consistent with these results.

Next we directly test our proposed interpretation of momentary crowding-out as taking a break, by having people engage in a brief unrelated task immediately after the reward period

[7] Indeed, in their review paper Deci et al. (1999) note that "studies with delayed assessments were virtually all with children" (p. 650).

ends, and then observing participants' behavior when they return to choosing between the target task and the videos.

STUDY 2: DIRECTLY TESTING THE EFFECT OF PROVIDING A 'BREAK' AFTER 'WORK'

Method

Online participants (N=257) were randomly assigned to one of six conditions. Two of the conditions (no-reward control; 5-cent per correct answer condition) were the same as in Study 1. In four "break" conditions, participants were given the 5 cent conditional incentive in Round 2, and then were asked to do an unrelated activity for 90 seconds immediately after the incentives ended. Ninety seconds was chosen to be equal to the duration of three trials, the approximate length of momentary crowding-out we observed in the prior studies.

As an additional test between our Effort-Balancing account and prior accounts, we tested two kinds of breaks. According to the Cognitive Evaluation theory, crowding-out occurs when people's sense of autonomy from a task has been undermined by external rewards. If this is the case, it is possible that by giving participants a series of explicit choices between the target task and the alternative leisure activity, we countered the autonomy-undermining effects of incentives, restoring intrinsic motivation and reducing the duration of post-reward crowding-out. Accordingly, we will compare breaks that involve choosing what to do during the break task (*choice-break*) to a version in which we assign the break task (*no-choice-break*). Giving choices has been found to bolster intrinsic motivation (Zuckerman et al., 1978; Moller, Deci, and Ryan, 2006;), so this additional opportunity to assert autonomy could eliminate momentary crowding-out in our setup if crowding-out is due to lack of autonomy.

However, if the break involving choices is instead perceived as effortful, it will not offset the effort exerted during the reward period, and therefore participants could still feel equally justified in choosing the leisure task after the break-period ends. Thus, if the *choice-break* is the perceived to be effortful, the Effort-Balancing account predicts momentary crowding-out after participants resume choosing between the two tasks. In contrast, if the *no-choice break* is perceived to be less effortful, the Effort-Balancing account suggests that it would be more likely to eliminate the justification and reduce momentary crowding-out. For generalizability, we will also test two different break tasks.

Participants in the break conditions were assigned to one of four conditions according to a 2 (Task Type: Writing, Logo Matching) x 2 (Choice: Yes, No) design. In the Writing task, participants wrote about a topic of general interest (e.g., why are tattoos so popular in today's society) whereas in the Logo Matching task participants matched brand-logos with brand-names for a particular product category (e.g., Oil and Gas companies). In the choice conditions, participants had to first choose between two topics or two product categories and then do the task. In the no-choice conditions, the topic or the product category was assigned to the participant. Each activity was of 30 seconds duration, and participants did three of the same type of activity for a total time of 90 seconds, after which the post-reward round resumed.

A pretest (N=56) confirmed that the choice conditions were seen to have provided significantly more sense of autonomy($p = .02$). However, in these conditions, participants had to go through and consider all the options, before doing the task. Participants perceived the choice tasks as more effortful ($p = .03$) and more work-like ($p = .002$). Thus, prior accounts would predict that the breaks involving choice would be more likely to reduce crowding-out, but our Effort-Balancing account predicts instead that the non-effortful breaks, that did not require

making additional choices, would offset effort expended, and therefore be more likely to reduce crowding-out.

Results

On average, participants chose to do the math task 66% of the time during the baseline period in the control condition, and there were no systematic changes over the trials. In the replication condition (5 cents per correct answer with no break, as in Study 1), immediately after the rewards ended there was a slight decrease in the choice of math tasks relative to control, consistent with momentary crowding-out (58% vs. 64%), but the difference was not significant in this study. The momentary crowding-out parameter was significant($\beta_{MCO} = -2.5, z = -2.6, p = .01$), but co-occurred with significant overall crowding-in ($\beta_{CO} = +1.1, z = +2.3, p = .02$).

In this study, we focus on the break conditions. The post-reward behavior in the two break conditions in which participants had to make break-related choices was similar, and therefore these two conditions were combined into a single *choice-break* condition. Likewise, the post-reward behavior in the conditions where participants were instead assigned to the break tasks was similar, and these were combined into a single *no-choice-break* activity condition.

When participants were given a post-reward break that involved making choices before starting the post-reward round (choice-break condition), the choice of the math task in the first trial of Round 3 was directionally lower in the choice-break condition (i.e., after a low-effort break), compared to the same trial in the control (48% vs. 63%, a 15 percentage-point difference; $t = 1.4, p = .18$). More importantly, the hierarchical regression model revealed significant momentary crowding-out relative to control ($\beta_{MCO} = -2.2, z = -2.1,\ p = .04$). Therefore, the

choice-break intervention failed to reduce post-reward crowding-out ($\beta_{MCO\ interaction} = +0.6, z = +0.8, p > .250$) relative to the replication condition.

However, even with the effortful break, post-reward crowding-out was momentary and after the initial dip, engagement did return to the baseline level with no long-term crowding-out of effort overall in Round 3 due to incentives (54% vs. 58%; $t < 1$). These findings were further confirmed in the regression model of Round 3 choices ($\beta_{CO} = +0.1, z = +0.3, p > .250$).

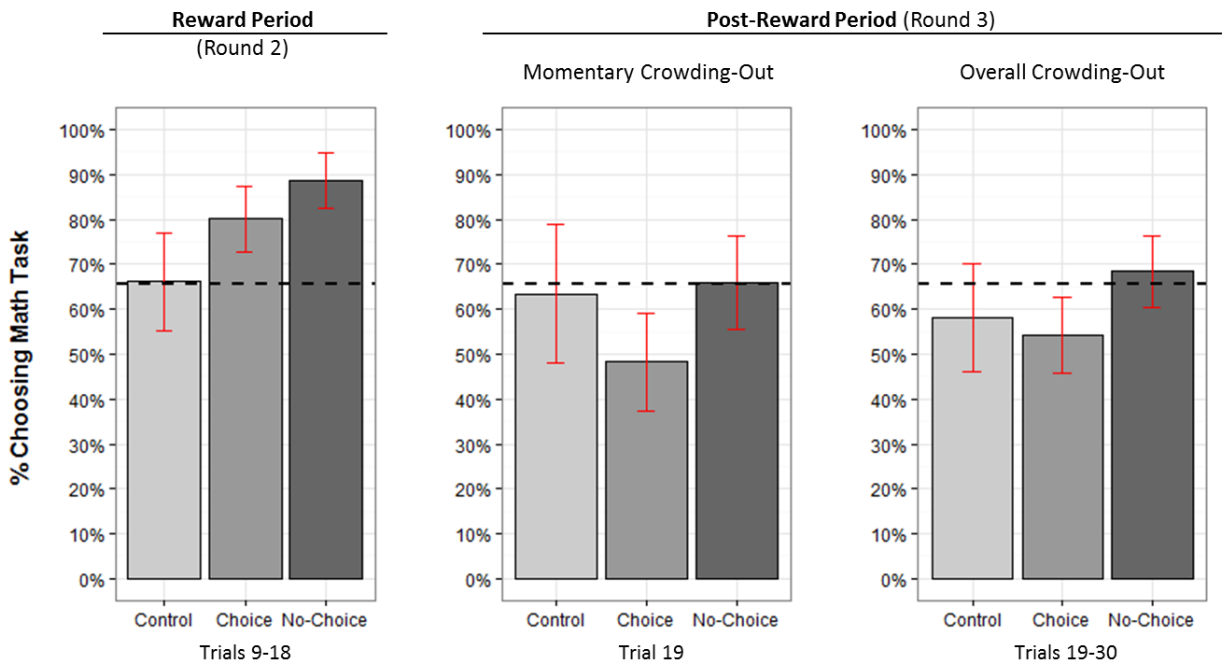

Figure 5: A no-choice, low-effort break after the incentive period eliminated momentary crowding-out (middle graph), but a more effortful choice-break does not. The dotted lines in the three graphs represent the average effort level across the conditions in the baseline (i.e., pre-incentive) period. The vertical lines are 95% CIs.

In contrast, when participants were given a post-reward break that did not involve making additional choices (no-choice-break condition), participants chose the math task 66% of the time in the next trial – similar to (and directionally higher than) the same trial in the control condition

(63%; $t < 1$; Figure 5). Further confirming that the low-effort break did not yield post-reward momentary crowding-out, we found no effect in the hierarchical regression model($\beta_{MCO} = -1.1, z = -1.0, p > .250$). Indeed, the no-choice-break intervention successfully reduced the momentary crowding-out effects observed in both the replication condition ($\beta_{MCO\ interaction} = +1.6, z = +2.1, p = .03$) as well as in the replication and choice-break conditions combined ($\beta_{MCO\ interaction} = +1.2, z = +1.9, p = .05$). Therefore, participants essentially returned to their baseline level of preference immediately after the low-effort break.

Furthermore, participants' engagement with the math task in the no-choice-break condition did not show any overall crowding-out compared to the average choice level in the control condition, but instead indicated directional crowding-in (68% vs. 58%; $t = 1.5, p = .15$). In fact, the slight increase in average proportion of choices of the math task represented a modest positive spillover effect of temporary incentives ($\beta_{CO} = +0.5, z = +2.0, p = .05$). This overall positive spillover was not significantly different from the replication condition ($\beta_{CO\ interaction} = -0.4, z = -0.9, p > .250$).

Therefore, giving people an autonomy-enhancing but effortful break (i.e., the choice-break) did not reduce post-reward momentary crowding-out, contrary to prior accounts. In fact, we observed significant momentary crowding-out after the choice-break. In contrast and consistent with our theory, after participants had a less-effortful break (i.e., in the no-choice-break condition), momentary crowding-out was eliminated. Thus, this study provides direct evidence for our characterization of post-reward crowding-out as people 'taking a break' after engaging in an incentivized activity, as long as the break does not require further effort (e.g., deliberating about options). This finding provides further evidence that momentary crowding-out is due to effort-balancing, rather than a reduction in autonomy.

The results of this study are also inconsistent with a simple 'forgetting' interpretation of momentary crowding-out, in which the salience of the conditional incentive is reduced after a few post-reward trials, returning participant motivation back to the pre-reward baseline level. Both break formats (with and without choices) were of the same total duration, and therefore would have facilitated forgetting about the incentives to the same extent, but only the low-effort break reduced momentary crowding-out.

In studies 1 and 2 we provided evidence for hypotheses 1a and 1b, which proposed that post-reward crowding-out is akin to taking a break after an effortful (yet interesting) task. In the next study, we test an important potential moderator of post-reward motivation that is predicted by our Effort-Balancing account the relative magnitude of the incentives. The prior accounts predict more crowding-out with higher incentives, since larger incentives would be experienced as more controlling and a stronger basis for self-perception inferences. Our Effort-Balancing account, however, predicts the opposite. If rewards are deemed to be sufficient to balance out the effort, people would find working on the task rewarding and would be less likely to feel that they need a break. Furthermore, larger rewards could also result in people persisting in the task at a higher level than their baseline preference, resulting in an overall positive spillover effect or post-reward crowding-in (Hypothesis 2).

STUDY 3: VARYING THE PERCEIVED MAGNITUDE OF REWARDS

Method

The participants (N=235) were randomly assigned to one of four conditions: two conditions which were identical to Study 1 (a no-reward control condition and a 5-cent

performance-contingent condition), and two other performance contingent-conditions in which we varied the relative reward amounts. In the *low-reward* condition, participants were told that they would earn 1 cent per correct answer in the second round, but were also told, as a basis for comparison, that in a previous version participants were paid 5 cents per correct answer. In the *high-reward* condition, participants were told they would earn 50 cents per correct answer, and that previous participants had earned 5 cents per correct answer. At the end of the study, participants were asked a few follow-up questions about their experience, including how much they liked the math task.

Results

On average, participants in the control condition chose the math task 60% of the time during Round 1 and there were no systematic trends across the trials. In the replication incentive condition (5 cents per correct answer), participants chose the math task 89% of the time when rewards were available, compared to 60% in the control condition during the same period (t=4.8, p<.001; Figure 6). The proportion in the replication condition fell to 45% in the first trial after the rewards ended, yielding marginally fewer choices of the math task than in the control condition during the same trial (56%; t=1.9, p=.05). However, crowding-out was momentary, and there was no difference in the average effort level between Round 3 and the baseline Round 1 in the replication and control conditions (60% vs. 53%; $t < 1$). More broadly, the hierarchical regression models found significant momentary crowding-out and no long-term effect of incentives ($\beta_{MCO} = -3.5, z = -3.1, p = .002$; $\beta_{CO} = +0.3, z = +0.4, p > .250$).

In the low-reward (1 cent per correct answer) condition, participants were marginally more likely to choose the math task during the rewarded trials in Round 2 than the control condition participants in the same trials (70% vs. 60%; t=1.9, p=.06). Therefore, even when the

relative framing highlighted the low incentive, effort increased while the incentive was available (unlike in Gneezy and Rustichini, 2000b), and no during-reward crowding-out for low incentives was observed in our experimental context.

However, immediately after the incentive ended in the low-reward condition, average math-task participation in the first trial of Round 3 was lower than in the control condition. Math task choices were directionally (11% percentage-points) lower in the first post-reward trial in the low-reward condition than in the same trial for the control condition (45% vs. 56%; $t = 1.47, p = .14$). As in the replication condition, the hierarchical regression model indicated significant momentary crowding-out ($\beta_{MCO} = -1.9\,, z = -2.0, p = .04$) in the low-reward condition. Once again, however, there was no long-term crowding-out (52% vs. 53%; $t < 1$) as was also confirmed in the regression model ($\beta_{CO} = -0.2, z = -0.3, p > .250$).

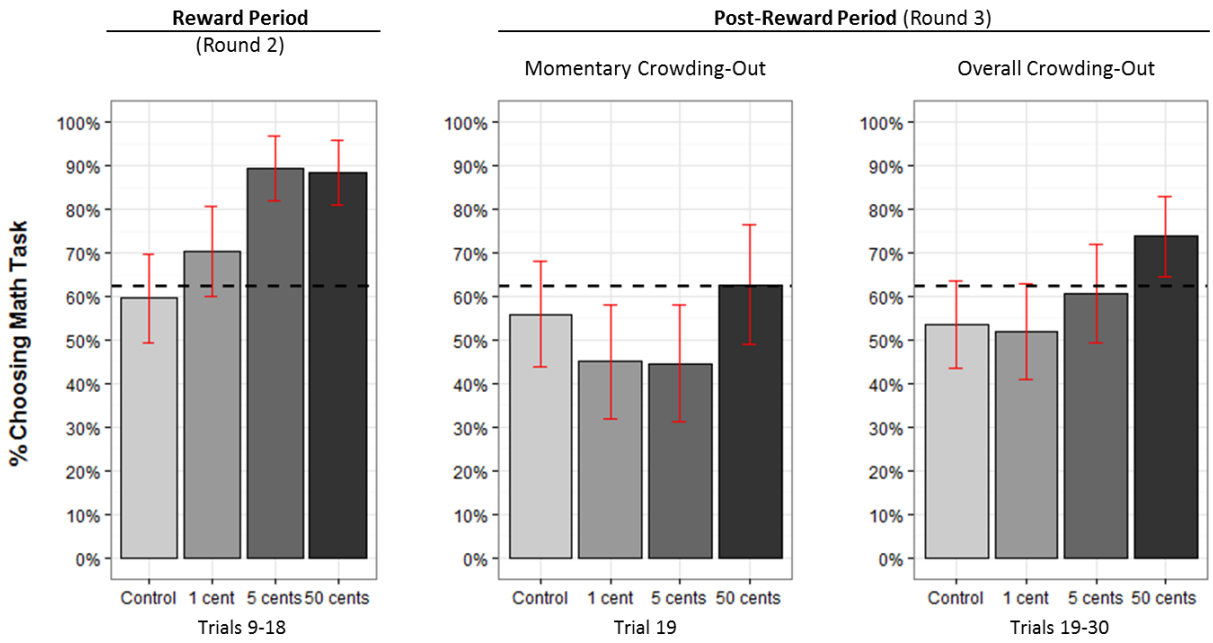

Figure 6: Average % math task choices in the control, replication (5 cents), low-reward (1 cent) and high-reward (50) conditions. The replication and low-reward conditions both exhibited momentary crowding-out, which was eliminated in the high-reward condition. The dotted lines in the three graphs represent the average effort level across the conditions in the baseline (i.e., pre-incentive) period. The vertical lines are 95% CIs.

Overall, post-reward behavior in Round 3 for the low-reward and replication conditions was very similar. Combining the replication and the low-reward conditions, we find significantly fewer math choices in the first trial after the incentives were stopped, compared to the same trial in the control condition (45% vs. 56%; $t = 1.98, p = .05$). These results were confirmed with the hierarchical regression model ($\beta_{MCO} = -2.7, z = -2.9, p = .003$). Moreover, there was no long-term crowding-out of effort in the low-reward condition (56% vs. 54%; $t < 1$) as was also confirmed in the regression model ($\beta_{CO} = +0.4, z = +0.07, p > .250$).

In contrast, people responded very differently when the incentive was substantially higher (50 cents per correct answer). The incentive in the high-reward condition (50 cents per correct answer) boosted effort while it was available, relative to the control condition (89% vs. 60%; t=5.5, p<.001), identical to the replication (5 cents) condition. The Effort-Balancing account predicts that, since the effort exerted during-reward was the same for the replication and high-reward conditions, the higher 50 cent incentive will lead to less post-reward crowding-out. Consistent with this prediction, the percentage of participants choosing the math task immediately after the high rewards were stopped was directionally *more* than for the same trial in the control condition (63% vs. 56%; t<1).

Overall, in the high-reward condition, participants did more of the math task during Round 3 compared to the control group in the same period (74% vs. 53%; $t = 3.7, p < .001$), reflecting a significant positive spillover effect of the larger incentive. This was also confirmed from the regression models ($\beta_{CO} = +1.9, z = +3.4, p < .001$)[8]. This suggests that participants did not feel the need for a break when they were earning 50 cents and the incentive was

[8] Because post-reward choices of the math task were higher in the high-reward condition than in control, the momentary crowding-out regression coefficient ($\beta_{MCO} = -1.6, z = -1.4, p = .2$) is not diagnostic, and instead measures the degree to which positive spillover was higher later in the round.

sufficiently rewarding to balance the effort, unlike the lower-reward conditions (1 cent and 5 cent), and they continued to favor the math task post-reward.

Furthermore, compared to both of the lower incentive (1 cent and 5 cents) conditions combined, people were significantly more likely to choose the math task in the trial immediately after the incentives were stopped when the incentive was high ($63\%\ vs. 45\%;\ t = 2.6, p = .009$). This further confirms that the momentary crowding-out behavior was significantly reduced (and, in fact, eliminated) in the high-reward condition. Results from hierarchical regression also confirmed that post-reward effort was significantly higher in the high-reward condition, compared to the low-reward and the replication conditions combined ($\beta_{CO\ interaction} = +2.2, z = +3.4, p < .001$). These findings are consistent with Effort-Balancing account, and the opposite of what prior autonomy-based or Overjustification accounts would have predicted.

Separating effort-balancing from other consequences of high effort

In general, an increase in task activity post-reward when incentives are higher could be explained by greater task exposure while incentivized, resulting in habit formation or reduced task difficulty due to having more practice. However, the effort expended during the incentive period was identical in the 50-cent and 5-cent conditions (both 89%). Contrary to the habit-formation or learning accounts, which would have predicted the same post-reward behavior, there was significantly more post-reward positive spillover in the high-reward condition than in the replication condition ($\beta_{CO\ interaction} = +1.9, z = +2.8, p = .004$).

In fact, consistent with theories of evaluative conditioning (Razran, 1954; De Houwer et al., 2001), paying high-rewards relative to the same effort (i.e., 50 cents) resulted in higher self-

reported liking for the math task at the end of the study, compared to the replication condition ($M_{50\ cents} = 6.9\ vs. M_{5\ cents} = 6.0;\ t(100) = 1.9, p = .05$). Indeed, participants rated working on the math tasks under high-rewards to be a significantly better opportunity for them compared to the replication condition ($M_{50\ cents} = 8.1\ vs. M_{5\ cents} = 6.5; t(100) = 4.9, p < .001$). This suggests that the post-reward lift in the high-reward condition is primarily driven by the effect on task motivation of the effort being more than balanced out by the large rewards.

The manipulation of relative incentive magnitude across conditions also allowed us to address potential alternative accounts. When incentives increase effort for the target task while the incentive is available (e.g., relative to control), this difference in experience with the task could directly affect subsequent behavior. If doing more of the task leads either to fatigue, or to depletion (Vohs et al., 2014) or to satiation with the task and need for variety, then when an incentive yields more-during reward task experience, the incentive could trigger a variety-seeking behavior that reduces task participation post-reward.

Participants did do significantly more of the math task in the 50-cent high-reward condition than in the 1-cent low-reward condition (88% vs 70%; $t(110) = 2.8, p = .006$), providing an opportunity to test these accounts. Contrary to the fatigue, depletion and satiation accounts, the larger number of math tasks done in Round 2 for the high incentive resulted in less, rather than more, subsequent momentary crowding-out. In fact, the high reward yielded overall crowding-in, relative to the low-reward condition($\beta_{CO\ interaction} = +2.4, z = +3.3, p < .001$). This is consistent with the Effort-Balancing account, particularly since participants subjectively evaluated the opportunity in the high-reward condition to be significantly better than the low-reward condition ($M_{50\ cents} = 8.1\ vs. M_{1\ cent} = 6.3;\ t(100) = 5.5, p < .001$).

Incentives can also create a reference point for the target task, such that after the end of the incentive, doing the target-task without pay is viewed as a relative loss, demotivating subsequent target-task activity. If this were the case, paying participants more should have also resulted in a higher reference point, and stronger crowding-out. Likewise, if the fact that we told participants up-front about the incentives ending was not sufficient to prevent disappointment effects, participants should have been more disappointed with the end of the larger incentive. The fact that crowding-out was eliminated, rather than strengthened, when the incentive was larger, suggests that reference points and disappointment do not explain our findings.

In the studies so far, the math task was framed as important and potentially beneficial, in part to highlight the self-control tradeoff between the work and leisure tasks. This raises the possibility, however, that the framing made participants feel obligated to work on the math task, rather than watch the videos, even after the reward ended. As a result, longer-term post-reward crowding-out of intrinsic motivation could have been reduced, reflecting participants' sense of obligation. In the next study, we address this potential concern.

STUDY 4: FRAMING BOTH CHOICE OPTIONS AS IMPORTANT

Method

Online participants (N=219) were randomly assigned to one of four conditions, in a 2 (Control, Reward) x 2 (Math Important, Both Important) between-subjects design. The two replication conditions (Math-Important control and reward) were similar to Study 1. The other two conditions (Both-Important control and reward) were the same, except that participants were told that their data, both from doing math and from the video task was important in the study.

Participants were also told that, since the survey was being administered to many people, it was completely up to them to choose what they wanted to do. This framing was designed to remove any signal to the participants that were expected to do the math tasks, and to encourage participants to choose what they truly wanted in do in each round. As a result, if participants' sense of obligation to do the math task had reduced the post-reward crowding-out in the prior studies, we would observe stronger crowding-out in the Both-Important condition.

Results

A manipulation check, collected at the end of the study, confirmed that participants in the Both-Important control condition expressed more agreement that the videos and math task were equally important (on a 9 point scale) than in the Math-Important control condition ($M_{control,both} = 4.4\ vs\ M_{control,math} = 3.7;\ t(110) = 2, p = .05$).

Since this study includes two differently-framed control conditions, we compare each reward condition to the corresponding control condition. We replicated the momentary crowding-out effect, when using the same instructions, in the Math-Important incentive and control conditions. Fewer people chose the math task in the first trial of Round 3 in the reward group after the incentives had ended, compared with the same trial in the control group (38% vs. 64%; t=2.1, p=.04; Figure 7). There was no long-term crowding-out of effort due to incentives, relative to control (57% vs. 63%; t <1). These results were further confirmed in the hierarchical regression models ($\beta_{MCO} = -4.7, z = -3.8, p < .001;\ \beta_{CO} = -0.04, z = -0.07, p > .250$).

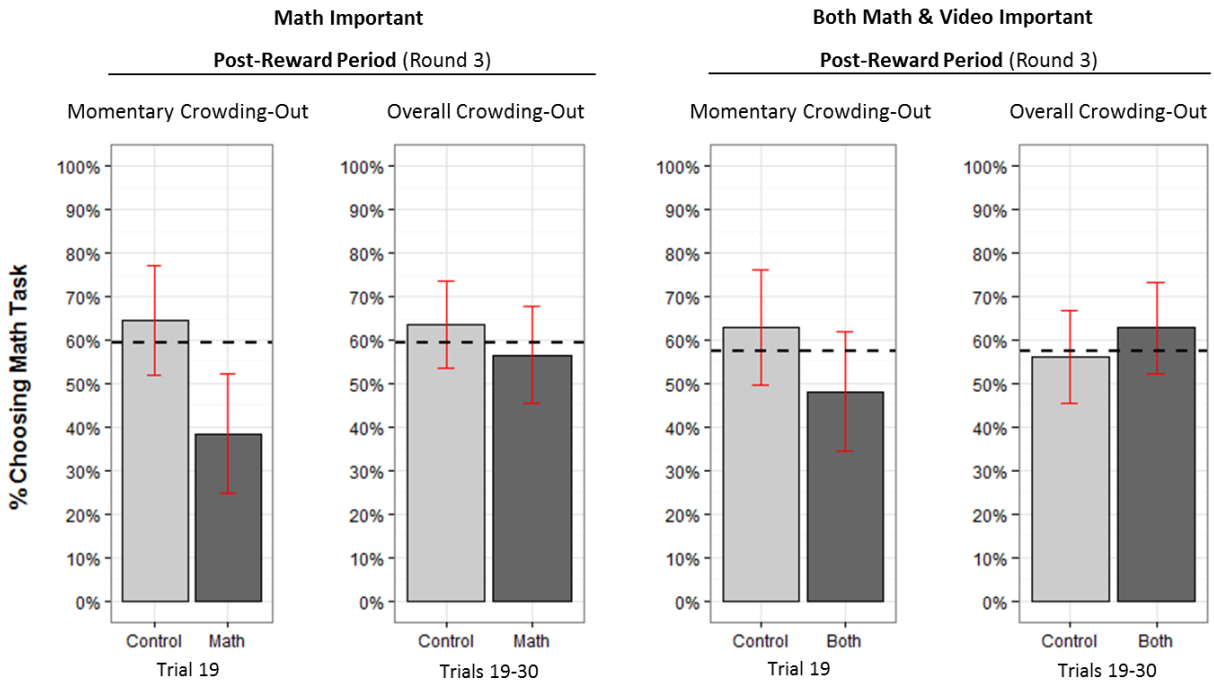

Figure 7: The left-panel of the figure shows the replication conditions, in which the math task was framed as important. The right-panel shows results when participants (separate control and reward group) were told that both math and video are equally important. The immediate post-reward results (trial 19) of these two conditions are very similar, and there is no evidence of overall crowding-out when both math and video were said to be equally important. The dotted lines in the three graphs represent the average effort level across the conditions in the baseline (i.e., pre-incentive) period. The vertical bars are 95% CIs.

Likewise, when we instead told participants that both the task options (math and video) are equally important, we again replicate the findings in the Both-Importance incentive and control conditions. Fewer people chose the math task in the first trial of Round 3 in the incentive condition after the incentives had ended, compared to in the control condition (48% vs. 63%; t=2.4, p=.02). There was no long-term crowding-out of effort due to incentives, relative to control (63% vs. 57%; t <1). These results were further confirmed in the hierarchical regressions ($\beta_{MCO} = -2.5, z = -3.3, p < .001; \beta_{CO} = +0.2, z = +0.5, p > .250$).

A hierarchical regression model also confirmed that there was no difference in the extent of momentary crowding-out between the two task-framing conditions ($\beta_{MCO\ interaction} =$

$+1.4, z = +1.08, p > .250$). Likewise, there was no difference in the extent of overall post-reward crowding-out between the two conditions ($\beta_{CO\ interaction} = +0.3, z = +0.4, p > .250$).

The results of this study suggest that the momentary nature of post-reward crowding-out is not explained by the experimental instructions inducing a feeling of obligation to do math tasks among the participants. Furthermore, the findings are also inconsistent with a signaling account, in which the participants continued with the more challenging math tasks (after a short break) to feel good about themselves.

In all the studies so far, we have incentivized participants for an effortful task, i.e., solving math problems. Incentives are generally deemed unnecessary to motivate people to do leisure tasks, and are therefore rarely applied to such tasks in practice. However, incentivizing the leisure task can provide a particularly useful test between the prior accounts and our proposed framework. According to the prior accounts, incentivizing a leisure task that is presumably intrinsically motivating should result in stronger crowding-out (Calder and Staw, 1975; Deci et al., 1999). However, the Effort-Balancing account would predict that, since doing the leisure task requires less effort for the same reward, the motive to take a break (and therefore momentary crowding-out) will not be stronger when incentivizing the less effortful leisure task, and could even be reduced, compared to the work task (Hypothesis 3).

STUDY 5: PAYING FOR A LEISURE TASK

Method

Online participants (N=340) were randomly assigned to one of four conditions in a 2 (Target task: Math vs. Video) x 2 (Control vs. Incentive) between-subjects design. In the two

reward groups, they were either paid for 5 cents for correctly completing math tasks, as in the prior studies, or paid 5 cents for each video they watched and rated (therefore, the video watching condition was different in all the cells of this study, compared to the previous studies in this paper, in that participants could rate the video after watching it). In the two control conditions, each matched one of the reward groups, highlighting the target-task, without any incentive[9].

Results

As in the previous study, we have two different control conditions, so we compare each reward condition to the corresponding control condition. We replicated the momentary crowding-out effect when math was the target task. Fewer people chose the math task in the first trial of Round 3 in the incentive condition after the rewards had ended, compared to in the control condition(47% vs. 68%; t=2.8, p=.006; Figure 8). There was no long-term crowding-out of effort due to incentives, relative to control (62% in both, t <1). These results were further confirmed in the hierarchical regression models ($\beta_{MCO} = -3.2, z = -4.6, p < .001$; $\beta_{CO} = +0.06, z = +0.1, p > .250$) .

[9] Two different framings of the math and video tasks were used in each condition, but since there were no differences, the results were merged. The details of these framings are in the supplemental materials.

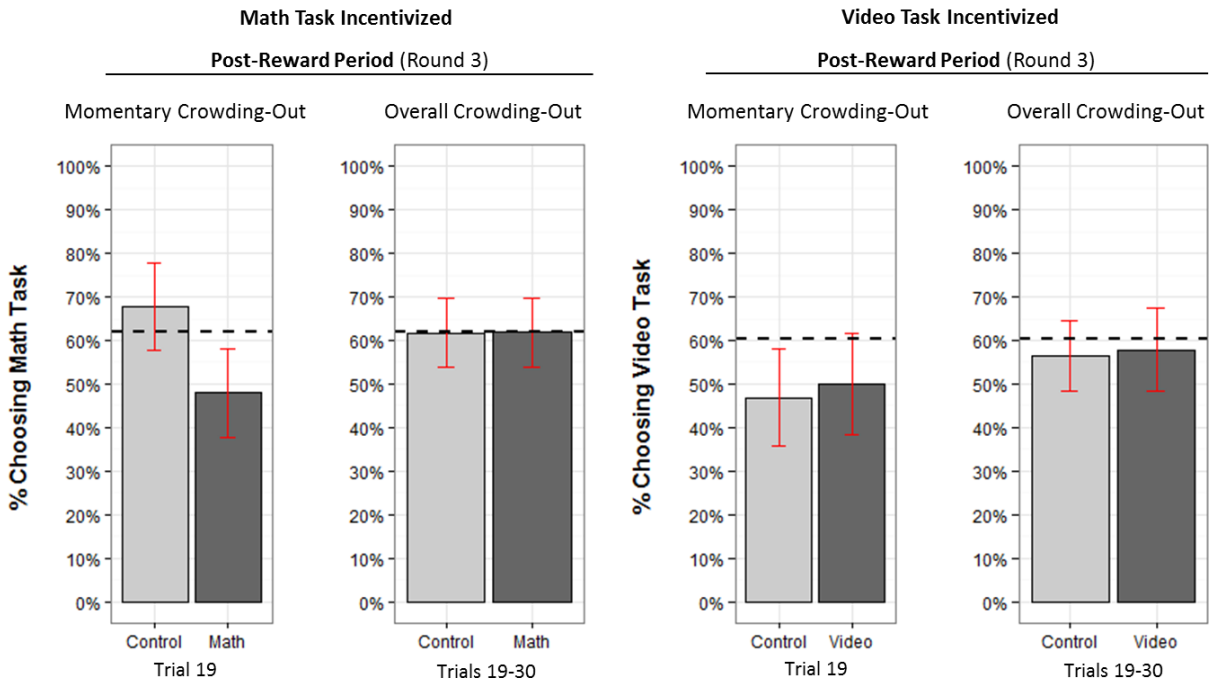

Figure 8: The left-panel of the figure shows the replication condition (control vs. 5 cents per correct math task). The right-panel shows no momentary crowding-out for the video-incentive conditions (control vs. 5 cents per rated video). The dotted lines in the three graphs represent the average effort level across the conditions in the baseline (i.e., pre-incentive) period. The vertical lines are 95% CIs.

However, when the video was the target task, participants chose the video task 50% of the time in the first trial of Round 3 compared to 47% of the time in the matching control condition. Therefore, when participants were paid for the leisure task, the proportion of video choices immediately after the rewards ended was actually directionally higher than the corresponding control condition. Thus, there was no evidence of momentary crowding-out when participants were paid for the leisure task ($t < 1$). There was also no long-term crowding-out of effort due to incentives, relative to control (57% vs. 56%; t <1). These results were further confirmed by the hierarchical regression models ($\beta_{MCO} = +0.1, z = +0.1, p > .250$; $\beta_{CO} = +0.1, z = +0.2, p > .250$).

Thus, contrary to the alternative accounts, neither momentary nor overall crowding-out were higher when a leisure task was incentivized. A hierarchical regression model confirmed that the momentary crowding-out was significantly reduced when videos were incentivized relative to when math tasks were incentivized($\beta_{MCO\ interaction} = +3.4, z = +2.9, p = .003$). However, beyond the difference in the immediate post-reward behavior after the end of the incentives, there was no difference between the two conditions in overall crowding-out ($\beta_{CO\ interaction} = +\ 0.03, z = \ +0.06, p > .250$).

Prior theories would predict more post-reward crowding-out when people are incentivized for a more entertaining leisure (vs. work) task. In contrast, we found that incentivizing people for the leisure task eliminated the post-reward crowding-out observed when offering the same incentives for the relatively more effortful math task. This finding is consistent with our proposed Effort-Balancing account. Since the leisure task involves less effort, and therefore requires little incentive to balance, reducing the justification to engage in the leisure activity decreases the motivation to take a break.

Task liking has often been used in the literature as a measure of intrinsic motivation (Deci et al., 1999), and the prior theories posited that rewards reduce intrinsic motivation for the incentivized task. Therefore, measures of task liking for an intrinsically motivating task would be predicted to decrease after such tasks were incentivized. In contrast, our account suggests that people will react more positively when the same incentive is offered for a less effortful task. As another test between these accounts, we measured liking of the tasks at the end of the study. When the video was incentivized, post-study liking for the video task was significantly higher than for the same task in the corresponding math-incentive condition $\left(M_{target=video} = 7.1\ vs. M_{target=math} = 6.2;\ \ t(159) = 2.9, p = .005\right)$. This suggests that providing an incentive

did not reduce intrinsic motivation for the video task, but instead made the low-effort leisure task seem even more rewarding.

GENERAL DISCUSSION

The Effort-Balancing account of post-incentive behavior

In this paper we proposed the Effort-Balancing account, a new way of thinking about the effects of incentives on post-reward effort for intrinsically motivating effortful tasks. We tested this account in five studies. Our findings suggest that, contrary to prior theories positing that incentives reduce a person's sense of autonomy from a task (Cognitive Evaluation Theory) or disrupt beliefs in one's own preferences (Overjustification theory), incentives may not reduce intrinsic motivation through these routes.

Instead, the increased effort induced by incentives can result in a desire to take a *break* when the incentive-period ends (Studies 1 and 2), in order to maintain a balance between effort and leisure. Thus, post-reward effort can decrease below the baseline level *momentarily*, but will return to baseline level when balance has been restored. Likewise, if people are given a non-effortful break after the incentivized period, their justification for a break will diminish and the momentary crowding-out will be reduced.

If the post-incentive break that yields momentary crowding-out is motivated by a desire to balance the effort, per our account, the effect should be moderated by the magnitude of the rewards. The larger the reward, the more it will balance out the effort, and the less people should

feel the need to take a break. When rewards were high, not only was momentary crowding-out reduced, but we even found a long-term positive spillover in engagement (Study 3), consistent with our account. Likewise, when a leisure task was incentivized, there was less effort that needed to be balanced, making it more difficult to justify taking a break, and momentary crowding-out was eliminated (Study 5). Contrary to the prior theories, our findings suggest that larger incentives and incentives for less effortful activities can have a *more* beneficial longer-term effect.

Our experimental paradigm enables us to control for several potential confounds that could also contribute to a post-reward decrease in engagement. Our findings cannot be explained by disappointment, since participants were informed from the very start of the reward-period about the temporary nature of the incentives. Likewise, the momentary crowding-out could not be explained by strategic behavior trying to induce a reintroduction of the rewards, since participants were informed that the incentives would only be available once in the entire experiment. Momentary crowding-out was also unlikely to reflect effort-shifting, since participants did not know the number of reward-trials and post-reward trials.

Participants were also unlikely to form broader inferences about the tasks based on characteristics of the incentives, since the pre-reward round gave participants enough opportunity to gain experience with the task. Furthermore, the results of testing several key moderators were inconsistent with additional potential alternative explanations, including participants forgetting about rewards with the passage of time (Study 2), depletion, satiation, variety-seeking and reference-points (Study 3), and social desirability or self-signaling (Study 4).

Our findings help to reconcile the disparate findings in the prior literature about the effect of temporary incentives on post-reward behavior. We do find immediate crowding-out when the

incentive ends, consistent with the robust effects in the lab-experiment literature, although our results support a different process explanation than previously assumed. In contrast, recent field studies have failed to find any detrimental effect of temporary incentives on post-reward engagement. These studies measured behavior quite some time after the incentives were stopped, and therefore may not have been able to detect potential momentary crowding-out.

Furthermore, many of these studies used substantial incentives, and so our framework also suggests that momentary crowding-out may not have occurred at all. Looking across the field studies, studies that provided larger incentives tended to find more positive long-term spillover effects (Kane, Johnson, Town, and Butler, 2004; Volpp et al., 2006, 2009), whereas studies with smaller incentives tended to find no spill-over, consistent with our account. Overall, our findings suggest that the psychological literature was correct about the existence of the effect, but the field study literature was correct about the scope of the effect, and our account can accommodate both sets of findings.

Implications for future research

Both our findings and the proposed Effort-Balancing account suggest important questions for future research. In our studies, we used a single paradigm repeatedly. It would be useful to test our framework using other tasks, as well as in other settings, especially in the field. In particular, it would be useful to examine the effect of temporary incentives in a longitudinal-setting where people are engaging in a task of their own volition, without a study-participation reward.

Although we find that the effect of temporary incentives on post-reward motivation is momentary, this could have more persistent consequences if a binding decision with long-term implications is made at the time when people want a break. Given the general finding that people often underestimate future adaptation (Gilbert, Pinel, Wilson, Blumberg, and Wheatley, 1998; Schkade and Kahneman, 1998), they may not anticipate that their motivation will return to baseline, and may make decisions (e.g., letting a gym membership expire) that reduce the likelihood of future re-engagement.

One common form of incentive is to provide a discount to motivate people to buy a product. There is a concern that offering such temporary promotions might result in crowding-out, with the probability of purchasing reduced below the baseline when the promotion ends (Dodson et al., 1978). However, it is not clear if even the momentary crowding-out due to Effort-Balancing would extend to situations where the effort takes the form of spending money. In such settings, it can be difficult to isolate the effect of decreased motivation from strategic shifts in purchase timing. However, our framework suggests that varying the discount magnitude could help distinguish between the two, since effort-balance crowding-out would be reduced for higher incentives (i.e., greater discounts), but time-shifting should be more prevalent for greater discounts.

Lastly, it would be useful to study the effect of temporary incentives on prosocial behavior. Pro-social tasks often convey a strong signaling-benefit when engaging in a task without any incentives. Past research has suggested that people might be less responsive to incentives for such activities, even when they are available (Dickinson, 1989; Benabou and Tirole, 2003; Ariely, Bracha, and Meier, 2009), and that introducing monetary rewards can undermine the effectiveness of pro-social incentives (Dube, Xueming, and Fang, 2015; Heyman

and Ariely, 2004; Yang, Urminsky, and Hsee, 2015). Therefore, it would be useful to examine if temporary monetary incentives result in only a momentary crowding-out of pro-social behaviors or a more persistent long-term crowding-out.

Re-opening the door to incentives

Incentives are one of the cornerstones of economic theory, and featured prominently in foundational theories developed in the early days of psychology as well. Rewards were considered to be a powerful reinforcer of desirable behavior (Skinner, 1953), an important determinant of motivational force in expectancy-valence theory (Vroom, 1964; Fishbein, 1967), and a means of creating and strengthening expectations of personal efficacy (Bandura, 1977). Despite the reinforcing nature of rewards, incentives are often viewed as counter-productive and subject to psychological backlash. Our results suggest that these concerns may have been over-stated, and that the ways in which people respond to incentives over time may involve different psychological mechanisms than previously thought. Our key insight in this paper is that post-reward crowding-out may often have more to do with people wanting a break after investing effort in their work, than with having their enthusiasm for the task "smothered" by incentives.

Taking a dynamic perspective on the effects of incentives in this research allowed us to uncover the momentary nature of crowding-out, and raises new questions about how incentives and motivation interact over time. What are the psychological factors that extend or inhibit momentary crowding-out? When will behaviors merely return to baseline after an incentive and when will positive spillover occur, increasing motivation long-term? What kinds of psychological interventions could be leveraged to make incentives more effective and motivating

in the long-term? We believe that our findings should not only re-open the door to testing the use of temporary incentives in consequential domains like health and education, but also to a new re-investigation of this most fundamental psychological driver of human motivation.

References

Ainslie, G. (1975). Specious reward: a behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, *82*(4), 463.

Ariely, D., Bracha, A., & Meier, S. (2009). Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *The American Economic Review*, 544–555.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191.

Benabou, R., & Tirole, J. (2003). Intrinsic and extrinsic motivation. *The Review of Economic Studies*, *70*(3), 489–520.

Bénabou, R., & Tirole, J. (2005). *Incentives and Prosocial Behavior* (Working Paper No. 11535). National Bureau of Economic Research.

Brooks, D. C., & Bouton, M. E. (1993). A retrieval cue for extinction attenuates spontaneous recovery. *Journal of Experimental Psychology: Animal Behavior Processes*, *19*(1), 77.

Brown, J. S. (1948). Gradients of approach and avoidance responses and their relation to level of motivation. *Journal of Comparative and Physiological Psychology*, *41*(6), 450.

Calder, B. J., & Staw, B. M. (1975). Self-perception of intrinsic and extrinsic motivation. *Journal of Personality and Social Psychology*, *31*(4), 599–605.

Charness, G., & Gneezy, U. (2009). Incentives to exercise. *Econometrica*, *77*(3), 909–931.

Chetty, R., Saez, E., & Sándor, L. (2014). *What Policies Increase Prosocial Behavior? An Experiment with Referees at the Journal of Public Economics*. National Bureau of Economic Research.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, *18*(1), 105–115.

Deci, E. L. (1972). The effects of contingent and noncontingent rewards and controls on intrinsic motivation. *Organizational Behavior and Human Performance*, *8*(2), 217–229.

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.

Deci, E. L., & Ryan, R. M. (1985). Cognitive evaluation theory. In *Intrinsic Motivation and Self-Determination in Human Behavior* (pp. 43–85). Springer.

De Houwer, J., Thomas, S., & Baeyens, F. (2001). Association learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin*, *127*(6), 853–869.

DelVecchio, D., Henard, D. H., & Freling, T. H. (2006). The effect of sales promotion on post-promotion brand preference: A meta-analysis. *Journal of Retailing*, *82*(3), 203–213. http://doi.org/10.1016/j.jretai.2005.10.001

Dickinson, A. M. (1989). The detrimental effects of extrinsic reinforcement on "intrinsic motivation." *The Behavior Analyst*, *12*(1), 1–15.

Dodson, J. A., Tybout, A. M., & Sternthal, B. (1978). Impact of deals and deal retraction on brand switching. *Journal of Marketing Research*, 72–81.

Dube, J. P., Xueming, L., & Fang, Z. (2015). Self-Signaling and Pro-Social Behavior: a cause marketing mobile field experiment. *Working Paper, University of Chicago*.

Esteves-Sorensen, C., Macera, R., & Broce, R. (2013). *Do Monetary Incentives Crowd Out Intrinsic Motivation? A Field Test in the Workplace*. mimeo.

Fehr, E., & Falk, A. (2002). Psychological foundations of incentives. *European Economic Review*, *46*(4), 687–724.

Fishbach, A., & Dhar, R. (2005). Goals as excuses or guides: The liberating effect of perceived goal progress on choice. *Journal of Consumer Research*, *32*(3), 370–377.

Fishbein, M. E. (1967). Readings in attitude theory and measurement.

Frey, B. S., & Jegen, R. (2001). Motivation crowding theory. *Journal of Economic Surveys*, *15*(5), 589–611.

Frey, B. S., & Oberholzer-Gee, F. (1997). The cost of price incentives: An empirical analysis of motivation crowding-out. *The American Economic Review*, 746–755.

Gilbert, D. T., Pinel, E. C., Wilson, T. D., Blumberg, S. J., & Wheatley, T. P. (1998). Immune neglect: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, *75*(3), 617.

Gneezy, U., Meier, S., & Rey-Biel, P. (2011). When and why incentives (don't) work to modify behavior. *The Journal of Economic Perspectives*, 191–209.

Gneezy, U., & Rustichini, A. (2000a). Fine is a price, a. *J. Legal Stud.*, *29*, 1.

Gneezy, U., & Rustichini, A. (2000b). Pay enough or don't pay at all. *The Quarterly Journal of Economics*, *115*(3), 791–810.

Halpern, S. D., French, B., Small, D. S., Saulsgiver, K., Harhay, M. O., Audrain-McGovern, J., … Volpp, K. G. (2015). Randomized trial of four financial-incentive programs for smoking cessation. *New England Journal of Medicine*.

Heyman, J., & Ariely, D. (2004). Effort for payment a tale of two markets. *Psychological Science*, *15*(11), 787–793.

Hoch, S. J., & Loewenstein, G. F. (1991). Time-Inconsistent Preferences and Consumer Self-Control. *Journal of Consumer Research*, *17*(4), 492–507.

Huber, J., & McCann, J. (1982). The impact of inferential beliefs on product evaluations. *Journal of Marketing Research*, 324–333.

Jackson, C. K. (2010). A Little Now for a Lot Later A Look at a Texas Advanced Placement Incentive Program. *Journal of Human Resources*, *45*(3), 591–639.

Jensen, M. C. (2003). Paying people to lie: The truth about the budgeting process. *European Financial Management*, *9*(3), 379–406.

Jhang, J. H., & Lynch, J. G. (2015). Pardon the Interruption: Goal Proximity, Perceived Spare Time, and Impatience. *Journal of Consumer Research*, *41*(5), 1267–1283.

John, L. K., Loewenstein, G., Troxel, A. B., Norton, L., Fassbender, J. E., & Volpp, K. G. (2011). Financial incentives for extended weight loss: a randomized, controlled trial. *Journal of General Internal Medicine*, *26*(6), 621–626.

Kalyanaram, G., & Winer, R. S. (1995). Empirical generalizations from reference price research. *Marketing Science*, *14*(3_supplement), G161–G169.

Kane, R. L., Johnson, P. E., Town, R. J., & Butler, M. (2004). A structured review of the effect of economic incentives on consumers' preventive behavior. *American Journal of Preventive Medicine*, *27*(4), 327–352.

Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist*, *28*(2), 107.

Khan, U., & Dhar, R. (2006). Licensing effect in consumer choice. *Journal of Marketing Research*, *43*(2), 259–266.

Kivetz, R. (2003). The Effects of Effort and Intrinsic Motivation on Risky Choice. *Marketing Science*, *22*(4), 477–502.

Kivetz, R., Urminsky, O., & Zheng, Y. (2006). The goal-gradient hypothesis resurrected: Purchase acceleration, illusionary goal progress, and customer retention. *Journal of Marketing Research*, 39–58.

Kivetz, R., & Zheng, Y. (2006). Determinants of justification and self-control. *Journal of Experimental Psychology: General*, *135*(4), 572.

Kohn, A. (1999). *Punished by rewards: The trouble with gold stars, incentive plans, A's, praise, and other bribes*. Houghton Mifflin Harcourt.

Kool, W., & Botvinick, M. (2014). A labor/leisure tradeoff in cognitive control. *Journal of Experimental Psychology: General*, *143*(1), 131.

Kruglanski, A. W., Alon, S., & Lewis, T. (1972). Retrospective misattribution and task enjoyment. *Journal of Experimental Social Psychology*, *8*(6), 493–501.

Kruglanski, A. W., Shah, J. Y., Fishbach, A., Friedman, R., Chun, W. Y., & Sleeth-Keppler, D. (2002). A theory of goal systems. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 34, pp. 331–378). Academic Press.

Lacetera, N., Macis, M., & Slonim, R. (2011). *Rewarding Altruism? A Natural Field Experiment* (Working Paper No. 17636). National Bureau of Economic Research.

Lacetera, N., Macis, M., & Slonim, R. (2013). Economic rewards to motivate blood donations. *Science*, *340*(6135), 927–928.

Lepper, M. R., & Greene, D. (1978). Overjustification research and beyond. *The Hidden Costs of Reward. Hillsdale, NJ: Erlbaurri*.

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.

Levin, I. P., Johnson, R. D., & Faraone, S. V. (1984). Information integration in price-quality tradeoffs: The effect of missing information. *Memory & Cognition*, *12*(1), 96–102.

Loewenstein, G., Brennan, T., & Volpp, K. G. (2007). Asymmetric paternalism to improve health behaviors. *Jama*, *298*(20), 2415–2417.

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research*, *45*(6), 633–644.

Mela, C. F., Jedidi, K., & Bowman, D. (1998). The long-term impact of promotions on consumer stockpiling behavior. *Journal of Marketing Research*, 250–262.

Moller, A. C., Deci, E. L., & Ryan, R. M. (2006). Choice and ego-depletion: The moderating role of autonomy. *Personality and Social Psychology Bulletin*, *32*(8), 1024.

Nisbett, R. E., & Valins, S. (1971). *Perceiving the causes of one's own behavior*. General Learning Press.

Oyer, P. (1998). Fiscal year ends and nonlinear incentive contracts: The effect on business seasonality. *Quarterly Journal of Economics*, 149–185.

Pink, D. H. (2011). *Drive: The Surprising Truth About What Motivates Us*. Penguin.

Pittman, T. S., Davey, M. E., Alafat, K. A., Wetherill, K. V., & Kramer, N. A. (1980). Informational versus controlling verbal rewards. *Personality and Social Psychology Bulletin*, *6*(2), 228–233.

Prendergast, C. (1999). The Provision of Incentives in Firms. *Journal of Economic Literature*, *37*(1), 7–63.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*. SAGE.

Razran, G. (1954). The conditioned evocation of attitudes (cognitive conditioning?). *Journal of Experimental Psychology*, *48*(4), 278.

Royer, H., Stehr, M. F., & Sydnor, J. R. (2012). *Incentives, Commitments and Habit Formation in Exercise: Evidence from a Field Experiment with Workers at a Fortune-500 Company*. National Bureau of Economic Research.

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.

Ryan, R. M., Mims, V., & Koestner, R. (1983). Relation of reward contingency and interpersonal context to intrinsic motivation: A review and test using cognitive evaluation theory. *Journal of Personality and Social Psychology*, *45*(4), 736.

Sandel, M. J. (2012). *What money can't buy: the moral limits of markets*. Macmillan.

Schkade, D. A., & Kahneman, D. (1998). Does living in California make people happy? A focusing illusion in judgments of life satisfaction. *Psychological Science*, *9*(5), 340.

Sela, A., Berger, J., & Liu, W. (2009). Variety, vice, and virtue: How assortment size influences option choice. *Journal of Consumer Research*, *35*(6), 941–951.

Shah, J. Y., & Kruglanski, A. W. (2002). Priming against your will: How accessible alternatives affect goal pursuit. *Journal of Experimental Social Psychology*, *38*(4), 368–383.

Shu, S. B., & Gneezy, A. (2010). Procrastination of Enjoyable Experiences. *Journal of Marketing Research*, *47*(5), 933–944.

Simonson, I., Carmon, Z., & O'curry, S. (1994). Experimental Evidence on the Negative Effect of Product Features and Sales Promotions on Brand Choice. *Marketing Science*, *13*(1), 23–40.

Skinner, B. F. (1953). *Science and human behavior*. Simon and Schuster.

Urminsky, O., & Goswami, I. (2015). Impatient to Achieve or Impatient to Receive: How the Goal Gradient Effect Underlies Time Discounting. *Workng Paper, University of Chicago*.

Urminsky, O., & Kivetz, R. (2011). Scope Insensitivity and the "Mere Token" Effect. *Journal of Marketing Research*, *48*(2), 282–295.

Urminsky, O., & Zauberman, G. (2015). The Psychology of Intertemporal Preferences. *Blackwell Handbook of Judgment and Decision Making, George Wu and Gideon Keren (eds), Wiley-Blackwell*.

Vohs, K. D., Baumeister, R. F., Schmeichel, B. J., Twenge, J. M., Nelson, N. M., & Tice, D. M. (2014). Making choices impairs subsequent self-control: a limited-resource account of decision making, self-regulation, and active initiative.

Volpp, K. G., Troxel, A. B., Pauly, M. V., Glick, H. A., Puig, A., Asch, D. A., … DeGuzman, J., others. (2009). A randomized, controlled trial of financial incentives for smoking cessation. *New England Journal of Medicine*, *360*(7), 699–709.

Volpp, K. G., Levy, A. G., Asch, D. A., Berlin, J. A., Murphy, J. J., Gomez, A., … Lerman, C. (2006). A randomized controlled trial of financial incentives for smoking cessation. *Cancer Epidemiology Biomarkers & Prevention*, *15*(1), 12–18.

Vroom, V. H. (1964). Work and motivation. 1964. *NY: John Wiley &sons*, 47–51.

White, R. W. (1959). Motivation reconsidered: the concept of competence. *Psychological Review*, *66*(5), 297.

Yang, A., Urminsky, O., & Hsee, C. K. (2015). Eager to Help yet Reluctant to Give: How Pro-social Effort and Pro-social Choices Diverge. *Working Paper, University of Chicago*.

Zuckerman, M., Porac, J., Lathin, D., & Deci, E. L. (1978). On the Importance of Self-Determination for Intrinsically-Motivated Behavior. *Personality and Social Psychology Bulletin*, *4*(3), 443–446.

Supplemental Materials at: http://home.uchicago.edu/~indranil/Papers/JMP_Apx_IndranilGoswami.pdf