frontiers in Psychology

Providing Extrinsic Reward for Test Performance Undermines Long-Term Memory Acquisition

Christof Kuhbandner[1*], Alp Aslan[2], Kathrin Emmerdinger[1], Kou Murayama[3]

[1]Psychology, University of Regensburg, Germany, Germany, [2]Psychology, University Martin-Luther-University Halle-Wittenberg, Germany, [3]Psychology, University of Reading, United Kingdom

Submitted to Journal:
Frontiers in Psychology

Specialty Section:
Educational Psychology

ISSN:
1664-1078

Article type:
Original Research Article

Received on:
13 May 2015

Accepted on:
13 Jan 2016

Provisional PDF published on:
13 Jan 2016

Frontiers website link:
www.frontiersin.org

Citation:
Kuhbandner C, Aslan A, Emmerdinger K and Murayama K(2016) Providing Extrinsic Reward for Test Performance Undermines Long-Term Memory Acquisition. *Front. Psychol.* 7:79. doi:10.3389/fpsyg.2016.00079

Providing Extrinsic Reward for Test Performance Undermines Long-Term Memory Acquisition

Christof Kuhbandner[1]*, Alp Aslan[2], Kathrin Emmerdinger[1], Kou Murayama[3]

[1]Department of Psychology, University of Regensburg, Regensburg, Germany
[2]Department of Psychology, University Martin-Luther-University Halle-Wittenberg, Halle, Germany
[3]Department of Psychology, University of Reading, Reading, United Kingdom

*** Correspondence**: Christof Kuhbandner, Department of Psychology, University of Regensburg, Universitaetsstr. 31, Regensburg, 93053, Germany.
christof.kuhbandner@ur.de

Keywords: testing effect, motivation, high-stakes testing, long-term memory, monetary reward.

Abstract

Based on numerous studies showing that testing studied material can improve long-term retention more than restudying the same material, it is often suggested that the number of tests in education should be increased to enhance knowledge acquisition. However, testing in real-life educational settings often entails a high degree of extrinsic motivation of learners due to the common practice of placing important consequences on the outcome of a test. Such an effect on the motivation of learners may undermine the beneficial effects of testing on long-term memory because it has been shown that extrinsic motivation can reduce the quality of learning. To examine this issue, participants learned foreign language vocabulary words, followed by an immediate test in which one third of the words were tested and one third restudied. To manipulate extrinsic motivation during immediate testing, participants received either monetary reward contingent on test performance or no reward. After one week, memory for all words was tested. In the immediate test, reward reduced correct recall and increased commission errors, indicating that reward reduced the number of items that can benefit from successful retrieval. The results in the delayed test revealed that reward additionally reduced the gain received from successful retrieval because memory for initially successfully retrieved words was lower in the reward condition. However, testing was still more effective than restudying under reward conditions because reward undermined long-term memory for concurrently restudied material as well. These findings indicate that providing performance-contingent reward in a test can undermine long-term knowledge acquisition.

1. Introduction

A central question of both experimental research and educational practice is how learning and retention can be promoted. A very powerful technique to improve long-term memory seems to be retrieving previously learned materials while taking a test (e.g., Gates, 1917; see Roediger & Butler, 2011, for a review). Several recent studies have renewed interest in this phenomenon by demonstrating that retrieving materials in a test promotes even better long-term retention than

restudying the same materials (e.g., Karpicke & Roediger, 2008; Roediger & Karpicke, 2006a), a phenomenon called "test-enhanced learning". In view of such findings, it has been recommended that the number of tests in education should be increased as frequent testing may boost students' achievement (Roediger & Karpicke, 2006b).

However, in real-life educational settings, test-taking may have additional effects on the emotions and motivations of learners, factors that have been largely neglected in previous research on the effect of testing. This neglect is particularly interesting because there is reason to assume that such effects may undermine the effectiveness of testing in enhancing long-term memory. For instance, regarding emotions, if a test induces a high degree of performance-related anxiety, the reduction in cognitive resources due to distraction by task-irrelevant emotion-induced thoughts (e.g., Ellis & Ashbrook, 1988) may impair cognitive processes underlying the enhancement of long-term learning. Indeed, this is supported by a recent study showing that performance-pressure induced test anxiety can attenuate the beneficial effects of a test on long-term memory (Hinze & Rapp, 2014).

At the motivational level, a typical effect of testing in real-life educational settings is that the motivation of learners is shifted towards an extrinsically motivated state due to the common practice of placing important consequences on the outcome of a test (for a review, see Harlen & Crick, 2003). Basically, whereas intrinsically motivated behaviors are engaged for their own sake, extrinsically motivated behaviors are driven by the prospect of instrumental gains and losses (e.g., Cerasoli, Nicklin, & Ford, 2014; Deci, 1971; Ryan & Deci, 2000). Critically, with regard to learning, numerous studies have shown that the quality of learning varies as a function of the motivational state of learners. Whereas intrinsically motivated learners show a more elaborative learning style characterized by more active and effortful learning that persists beyond the point of being rewarded or punished, extrinsically motivated learners show a more superficial learning style characterized by more passive and less effortful learning that vanishes beyond the point of being rewarded or punished (e.g., Benware & Deci, 1984; Lepper, Corpus, & Iyengar, 2005, Dewitte, & Lens, 2004; Vansteenkiste, Simons, Lens, Sheldon, & Deci, 2004). Accordingly, it may be that when the taking of a test leads to a high degree of extrinsic motivation, the detrimental effects of extrinsic motivation on learning may undermine the memory-enhancing effect of retrieving learned material in a test.

Basically, there are two possibilities why a test that induces a high degree of extrinsic motivation may undermine the effectiveness of testing in enhancing long-term knowledge acquisition. First, by providing gains contingent on performance, such a test induces a strong desire to perform as well as possible. Such a desire may impair the quality of retrieval of actually stored knowledge. On the one hand, the rate of successfully retrieved information may be decreased because it has been shown that people often perform below actual abilities when trying to perform as well as possible, an observation that is commonly attributed to the experience of performance pressure. Such performance pressure often leads to the occupation of attention by task-irrelevant thoughts, such as ruminations about one's performance and its consequences (e.g., Baumeister, 1984; DeCaro, Thomas, Albert, & Beilock, 2011). On the other hand, the rate of erroneously retrieved information (i.e., commission errors) may be increased because people may try to maximize their gains by guessing (e.g., Legault, & Inzlicht, 2013). An increased rate of commission errors in a test may be problematic for long-term learning because learners may store the erroneously retrieved information in long-term memory, with the detrimental consequence that they may acquire erroneous knowledge (e.g., Roediger & Karpicke, 2006). In fact, the reasonableness of such an assumption is supported by the observation in the present study that in the delayed test, commission errors for tested vocabulary pairs were observed more often in the high than the low extrinsic-motivation condition.

These detrimental effects of extrinsic reward on the retrieval of learned material in a test may decrease the benefits gained from testing for long-term memory. Second, a test that induces extrinsic motivation may even reduce the benefit received from successful retrieval. Most theoretical accounts proposed to explain the high effectiveness of testing assume that retrieval of information from memory represents a new learning event (i.e., reconsolidation; e.g., Dudai, 2004) that allows storing the retrieved information more elaborately and deeply (e.g., Finn & Roediger, 2011; Finn, Roediger, & Rosenzweig, 2012; see Roediger & Butler, 2011, for a review). However, if a test is taken in an extrinsically motivated state, such reconsolidation processes may be weakened due to the more passive and less persistent learning brought about by extrinsic motivation (e.g., Benware & Deci, 1984; Grolnick & Ryan, 1987; Vansteenkiste et al., 2004).

The aim of the present study was to examine the effect of extrinsic motivation on the long-term memory effects of testing. To examine the issue, we employed a standard testing-effect paradigm and manipulated the degree of extrinsic motivation during immediate testing. Participants first studied Swahili–German vocabulary pairs (e.g., Mashua–Boat) without mentioning that they may be rewarded for their later test performance. In a subsequent immediate memory test, one third of the vocabulary pairs were tested, one third were presented for restudy, and the remaining third did not appear in the test and served as control pairs. In order to manipulate the degree of extrinsic motivation during immediate testing, participants received either performance-contingent monetary reward for test performance (high extrinsic-motivation condition), or not (low extrinsic-motivation condition; e.g., Murayama, Matsumoto, Izuma, & Matsumoto, 2010). To control for potential confounding effects of receiving money on post-learning consolidation processes (e.g., Murayama & Kitagami, 2014; Nielson & Bryant, 2005), participants in the low extrinsic-motivation condition received money as well. However, other than in the high extrinsic-motivation condition, this money was not performance-contingent and not framed as a reward. Instead, participants took part in a lottery, and they were told that they can earn some additional remuneration for participating in the experiment. Then, after a delay of one week, memory for all initially studied vocabulary pairs was tested.

In the immediate test, we expected to replicate the detrimental effects of providing extrinsic reward contingent on test performance on the quality of retrieval of learned knowledge (e.g., Baumeister, 1984; DeCaro et al., 2011; Legault, & Inzlicht, 2013); that is, we expected that the rate of successfully retrieved information would be decreased and the rate of commission errors would be increased. If so, then memory in the delayed test for initially tested vocabulary pairs should be impaired in the high compared to the low extrinsic-motivation condition as well because less vocabulary pairs can benefit from being initially successfully retrieved. If extrinsic motivation additionally undermines the benefit received from successful retrieval, then memory for initially successfully retrieved vocabulary pairs in the delayed test should be reduced in the high compared to the low extrinsic-motivation condition as well. Regarding the effect of extrinsic motivation on restudied items, it may be that concurrently restudied items suffer less from extrinsic motivation because the problem of retrieval impairment is circumvented when all information is presented again for restudy. If so, the advantage of testing over restudying should be decreased in the high compared to the low extrinsic-motivation conditions. However, extrinsic motivation may lead to a less effortful restudying of concurrently presented but not rewarded information. If so, despite the detrimental effect of extrinsic motivation on the effects of testing, the advantage of testing over restudying should be similar between the high and low extrinsic-motivation conditions.

2. Materials and Methods

2.1. Participants

Sixty undergraduate students (49 females; *M* = 22.9, *SD* = 4.3 years) participated in the experiment for course credit. Participants were tested in small groups of up to five individuals. One of the original participants was replaced (in the low extrinsic-motivation condition) because he did not recall a single item in the immediate test. Including this participant did not change the significance of any of our results. The study was conducted in accordance with the Helsinki declaration and the University Research Ethics Standards.

2.2. Materials

The study list consisted of 30 Swahili-German vocabulary pairs drawn from Karpicke and Roediger (2008).

2.3. Design and Procedure

Half of the participants were randomly assigned to a low extrinsic-motivation condition, and the other half to a high extrinsic-motivation condition. In each condition, the experiment consisted of three main phases: a study phase, a (combined) test/restudy phase, and a one-week delayed final test phase. In the study phase, participants were presented 30 vocabulary pairs (e.g., *Mashua* – Boat, *Bustani* – Garden, *Farasi* – Horse) in randomized order. Stimuli were delivered via a projector at a 7-sec rate with an interstimulus interval of 1 sec. Participants were asked to read the vocabulary pairs silently and memorize them for a later cued-recall test (e.g., *Mashua* – ?). It was not mentioned that they may be rewarded for their later test performance. Following presentation of the last pair, the whole list was presented a second time. In the subsequent test/restudy phase, participants were tested on one third of the vocabulary pairs (without any feedback) by providing the Swahili words as retrieval cues for the German words (tested vocabulary pairs; e.g., *Mashua* – ?), while another third of the pairs were re-presented to the participants for restudying (restudied vocabulary pairs; e.g., *Bustani* – Garden); the remaining third of vocabulary pairs did not appear in this phase and served as a baseline for the benefits gained from testing and restudying (control vocabulary pairs). The stimuli were delivered via a projector and participants were instructed to write down both of the two words of a vocabulary pair within 10 sec, both for the test and restudy pairs. The order of the 10 test and 10 restudy trials was randomized, and the assignment of the vocabulary pairs to the three learning conditions was counterbalanced. Directly before the test/restudy phase, participants in the high extrinsic-motivation condition were encouraged to perform as well as possible on the test trials because they were told they would be paid 1 Euro for each correctly recalled German word. No such instruction was given in the low extrinsic-motivation condition. In order to control for potential confounding effects of receiving money on post-learning consolidation processes (e.g., Murayama & Kitagami, 2014; Nielson & Bryant, 2005), participants in the low extrinsic-motivation condition also received money. In contrast to the high extrinsic-motivation condition, however, this remuneration was not related to their recall performance and was not framed as 'reward'. Instead, participants were told that they can take part in a lottery where they could earn some additional remuneration for participating in the experiment (additionally to the course credit they received for participation). The sums of money that individual participants won in the lottery were adjusted so that each participant in the low extrinsic-motivation condition was monetarily yoked to a participant in the high extrinsic-motivation condition so that, across participants, the mean amount of received money was equal in the two conditions. After immediate testing, all participants were asked to return to the laboratory one week later for a delayed cued-recall test covering all initially studied vocabulary pairs, and they were informed that the delayed memory test would be unpaid. Upon arrival in the laboratory one week later, participants were given a sheet of paper including the 30 Swahili words in randomized

order, and were asked to recall and write down the corresponding German words. There was no time restriction in this test. After completion of the delayed memory test, participants were thanked and they received their money.

3. Results

3.1. Immediate Test

Memory performance in the immediate test as a function of motivational condition is shown in Figure 1A. Probability of correct recall was lower in the high than the low extrinsic-motivation condition (high: $M = 0.60$, $SD = 0.24$ vs. low: $M = 0.74$, $SD = 0.21$), $t(58) = -2.47$, $p = .017$, $d = 0.64$, whereas the probability of commission errors (intralist intrusions) was higher in the high than the low extrinsic-motivation condition (high: $M = 0.10$, $SD = 0.12$ vs. low: $M = 0.02$, $SD = 0.04$), $t(58) = 3.51$, $p < .001$, $d = 0.91$.

3.2. Delayed Test

Figure 1B shows memory performance in the delayed test for initially tested, restudied, and control vocabulary pairs as a function of motivational condition. A 2 x 3 ANOVA with the within-participants factor of vocabulary type (tested, restudied, control) and the between-participants factor of extrinsic motivation (high, low) revealed a significant main effect of vocabulary type, $F(2, 116) = 47.48$, $p < .001$, $\eta_p^2 = .450$, reflecting the fact that correct recall was higher for restudied than for control vocabulary pairs ($M = 0.33$, $SD = 0.25$ vs. $M = 0.18$, $SD = 0.18$), $t(59) = 5.60$, $p < .001$, $d = 0.72$, and even higher for tested than for restudied vocabulary pairs ($M = 0.46$, $SD = 0.27$), $t(58) = 4.15$, $p < .001$, $d = 0.48$. The main effect of extrinsic motivation was also significant, $F(1, 58) = 7.40$, $p = .009$, $\eta_p^2 = .113$, reflecting the fact that, collapsed across the three vocabulary types, correct recall was lower in the high than the low extrinsic-motivation condition (high: $M = 0.26$, $SD = 0.17$ vs. low: $M = 0.39$, $SD = 0.21$). The interaction between vocabulary type and extrinsic motivation was not significant, $F(2, 116) = 1.92$, $p = .151$, $\eta_p^2 = .032$. Simple main effect analyses showed that correct recall for tested vocabulary pairs was lower in the high than the low extrinsic-motivation condition (high: $M = 0.38$, $SD = 0.28$ vs. low: $M = 0.54$, $SD = 0.24$), $t(58) = -2.43$, $p = .018$, $d = 0.63$, indicating that initial reward reduced memory performance for tested contents. Correct recall for restudied vocabulary pairs was lower in the high than the low extrinsic-motivation condition as well (high: $M = 0.25$, $SD = 0.19$ vs. low: $M = 0.42$, $SD = 0.28$), $t(58) = -2.71$, $p = .009$, $d = 0.70$, indicating that the detrimental effects of reward transferred to restudied items. Correct recall for control vocabulary pairs did not significantly differ between motivational conditions (high: $M = 0.14$, $SD = 0.14$ vs. low: $M = 0.21$, $SD = 0.21$), $t(58) = -1.45$, $p = .153$, $d = 0.37$.

A 2 x 3 ANOVA with the within-participants factor of vocabulary type (tested, restudied, control) and the between-participants factor of extrinsic motivation (high, low) on the probability of commission errors (intralist intrusions) revealed neither a main effect of vocabulary type nor a main effect of extrinsic motivation, Fs < 2.48, ps > .121, but a significant interaction, $F(1, 58) = 3.33$, $p = .039$, $\eta_p^2 = .054$. Simple main effect analyses showed that whereas commission errors did not differ between motivational conditions for restudied (high: $M = 0.03$, $SD = 0.07$ vs. low: $M = 0.02$, $SD = 0.05$) and control vocabulary pairs (high: $M = 0.03$, $SD = 0.05$ vs. low: M = 0.043, $SD = 0.06$), ts < 0.70, ps > .490, for tested vocabulary pairs commission errors were observed more often in the high than the low extrinsic-motivation condition (high: $M = 0.07$, $SD = 0.09$ vs. low: $M = 0.020$, $SD = 0.05$), $t(58) = 2.54$, $p = .014$, $d = 0.66$.

Finally, we examined the effect of reward on memory for vocabulary pairs which were initially successfully retrieved. To control for potential item-selection artifacts (i.e., artifacts due to unbalanced distribution of vocabulary pairs across conditions because of differential recall in the immediate memory test), we determined for each vocabulary pair the conditional probability of correct recall in the delayed test given successful recall in the immediate test, collapsing data across participants. As shown in Figure 1C, conditional probability of correct recall was lower in the high than the low extrinsic-motivation condition (high: $M = 0.58$, $SD = 0.26$ vs. low: $M = 0.70$, $SD = 0.19$), $t(29) = 2.70$, $p = .011$, $d = 0.49$, indicating that even initially successfully retrieved vocabulary pairs benefited less from testing when extrinsic motivation was high.

4. Discussion

Previous research has shown that retrieving previously learned contents in a test can improve long-term memory for tested contents, suggesting that the number of tests in education should be increased to enhance knowledge acquisition (see Roediger & Karpicke, 2006, for a review). The present study demonstrates, however, that the effect of retrieval is undermined when a test entails a high degree of extrinsic motivation due to the provision of gains contingent on test performance. Compared to a no-reward condition, rewarding participants with money depending on performance in the immediate test decreased correct recall and increased commission errors for tested contents after one week. Thus, given that the placing of important consequences on the outcome of a test is common practice in educational settings, the consequences of testing in education on the acquisition of knowledge for later life and work may be less encouraging than previously believed.

More detailed analyses showed that the detrimental effects of reward were attributable to two factors. First, the provision of monetary reward contingent on performance reduced correct recall and increased commission errors in the immediate test, a pattern that typically occurs in situations where people try to perform as well as possible to maximize promised extrinsic gains (e.g., Baumeister, 1984, DeCaro et al., 2011; Legault, & Inzlicht, 2013). Such an effect of reward on immediate test performance seems to have two negative consequences for later long-term memory. On the one hand, by decreasing the amount of information that is successfully retrieved, reward seems to reduce the amount of stored information that can benefit from retrieval (e.g., Bjork & Bjork, 1992; Kornell, Bjork, & Garcia, 2011). On the other hand, by increasing the amount of information that is erroneously retrieved, reward seems to increase the degree of information that is erroneously reconsolidated. This is reflected by the fact that commission errors in the delayed test were increased in the reward condition for vocabulary pairs that were part of the immediate memory test, but not for restudied and control vocabulary pairs that were not actively retrieved during immediate testing.

Second, even for vocabulary pairs that were initially successfully retrieved, long-term memory was reduced when reward was initially provided. Thus, extrinsic motivation seems to undermine even the benefit gained from successfully retrieving stored information in a test. Such an effect is consistent with findings showing that the quality of learning differs depending on motivational state. Compared to intrinsic motivation, extrinsically motivated learners show a less elaborative learning style characterized by more passive and less effortful learning that vanishes beyond the point of being rewarded or punished (e.g., Benware & Deci, 1984; Grolnick & Ryan, 1987; Lepper et al., 2005; Vansteenkiste et al., 2004). Such detrimental effects of extrinsic motivation may reduce the memory-enhancing effects of testing by reducing the quality of learning evoked by retrieval.

The present results further show that high extrinsic motivation can even have detrimental effects on long-term memory for material that is restudied. Compared to the no-reward condition, memory in

the delayed test for vocabulary pairs that were initially restudied was reduced in the reward condition as well. As participants were forced to write down each of the to-be-restudied vocabulary pairs during restudy, such a finding cannot easily be explained by the simple assumption that rewarding participants only for some vocabulary pairs led them to abandon processing of not rewarded vocabulary pairs. However, in such a situation, the induced extrinsic motivation seems to bring about a less effortful restudying of not rewarded information.

In the present study, we examined the effect of providing performance-contingent reward in an immediate memory test on performance in a delayed long-term memory test where no reward was provided. This situation mimics the typical educational scenario in which the objective is to provide learners with knowledge to prepare them for later life and work, where knowledge retrieval is not necessarily driven by extrinsic forces. Doing so, we found that providing extrinsic reward for test performance can undermine long-term knowledge acquisition of the assessed contents. The situation may be different however, when extrinsic motivation is increased during immediate test-taking because learners are aware that they are preparing for a delayed test for which they will be rewarded based on their performance. In such a situation, additional motivational factors may play an important role during immediate test taking, such as the motivation to learn the material as well as possible for the delayed test (see Hidi, & Harackiewicz, 2000, for a review). As a result, this may attenuate the detrimental effects of extrinsic motivation on the quality of learning. Indeed, this assumption is supported by a recent study, showing that the prospect of receiving monetary reward for performance in the delayed test seems not to reduce the beneficial effects testing (Kang & Pashler, 2014). Still, it seems possible that long-term knowledge acquisition beyond the delayed test for which reward was provided suffers from the increase in extrinsic motivation, which is an issue that should be explored in future research.

The present study also raises several questions that should be addressed in future research. First, our sample consisted mainly of female undergraduate students. Thus, future research should examine whether the results of the present study generalize across gender and different levels of education. Second, in order to be able to relate our results to prior findings, the study material consisted of foreign language word pairs that have been frequently used in research on the effects of testing (e.g., Kang & Pashler, 2014; Karpicke & Roediger, 2008). Thus, future research should examine whether the results of the present study generalize across other types of study materials such as text passages or general knowledge facts. Third, because all participants participated for course credit, learning in the condition where no reward was provided for test performance was not entirely intrinsically motivated. Our prediction would be that the detrimental effect of reward may be even more noticeable when compared to a condition where participants participate without receiving any reward because their motivational state is then shifted even more strongly towards an intrinsically motivated state, a prediction that deserves future research.

Finally, the present results may have important implications for applied settings, such as educational practice. Based on the finding that test taking can enhance later memory, it has been argued that increasing the number of tests in education is a promising technique to boost educational achievement (e.g., Roediger & Karpicke, 2006a). Our findings demonstrate that the effectiveness of testing in improving long-term knowledge acquisition is reduced when a test leads to a high degree of extrinsic motivation due to the provision of performance-contingent reward. Therefore, the common practice to implement tests as high-stakes assessments which have to be passed in order to reach important benefits may counteract the beneficial effects of testing on the acquisition of knowledge in long-term memory. One possibility to at least partly overcome the detrimental effects of reward may be to provide corrective feedback as this would reduce the problem of retrieval impairment due to the

desire to perform as well as possible. However, as extrinsic motivation even seems to decrease the memory strength gained from successful retrieval and from restudying concurrently presented contents, tests that lead to high extrinsic motivation may still be less effective than tests that do not increase extrinsic motivation. Thus, if possible, educators would be well advised to implement tests as low-stakes assessments, in order to maximize the effectiveness of testing for long-term knowledge acquisition.

Author Contributions

CK developed the study concept. All authors contributed to the study design and all authors analyzed and interpreted the data. CK prepared the draft manuscript, and AA, KE, and KM provided critical revisions. All authors approved the final version of the manuscript for submission.

References

Baumeister, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *J. Pers. Soc. Psychol.* 46, 610–620. doi:10.1037//0022-3514.46.3.610

Benware, C., and Deci, E. L. (1984). The quality of learning with an active versus passive motivational set. *Am. Educ. Res. J.* 21, 755–765. doi:10.3102/00028312021004755

Bjork, R. A., and Bjork, E. L. (1992). "A new theory of disuse and an old theory of stimulus fluctuation," in *From learning processes to cognitive processes: Essays in honor of William K. Estes (Vol. 2)*, eds. A. Healy, S. Kosslyn, and R. Shiffrin (Hillsdale, NJ: Erlbaum), 35-67.

Cerasoli, C. P., Nicklin, J., and Ford, M. T. (2014). Intrinsic motivation and extrinsic incentives jointly predict performance: A 40-year meta-analysis. *Psychol. Bull.* 140, 980–1008. doi:10.1037/a0035661

DeCaro, M. S., Thomas, R. D., Albert, N. B., and Beilock, S. L. (2011). Choking under pressure: Multiple routes to skill failure. *J. Exp. Psychol. Gen.* 140, 390–406. doi:10.1037/a0023466

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *J. Pers. Soc. Psychol.* 18, 105–115. doi:10.1037/h0030644

Dudai, Y. (2004). The neurobiology of consolidations, or, how stable is the engram?. *Annu. Rev. Psychol.* 55, 51–86. doi:10.1146/annurev.psych.55.090902.142050

Ellis, H. C., and Ashbrook, P. W. (1988). "Resource allocation model of the effects of depressed mood states on memory", in *Affect, cognition, and social behaviour*, eds. K. Fiedler and J. Forgas (Toronto: Hogrefe), 25-32.

Finn, B., and Roediger, H. L. (2011). Enhancing retention through reconsolidation: Negative emotional arousal following retrieval enhances later recall. *Psychol. Sci.* 22, 781-786. doi:10.1177/0956797611407932

Finn, B., Roediger, H. L., and Rosenzweig, E. (2012). Reconsolidation from negative emotional pictures: Is successful retrieval required? *Mem. Cognition* 40, 1031-1045. doi:10.3758/s13421-012-0203-7

Gates, A. I. (1917). Recitation as a factor in memorizing. *Arch. Psychol.* 6(40).

Grolnick, W. S., and Ryan, R. M. (1987). Autonomy in children's learning: An experimental and individual difference investigation. *J. Pers. Soc. Psychol.* 52, 890-898. doi:10.1037/0022-3514.52.5.890

Harlen, W., and Crick, R. D. (2003). Testing and motivation for learning. *Assessment Educ.* 10, 169–207. doi:10.1080/0969594032000121270

Hidi, S., and Harackiewicz, J. M. (2000). Motivating the academically unmotivated: A critical issue for the 21st century. *Rev. Educ. Res.* 70, 151–179. doi:10.3102/00346543070002151

Hinze, S. R., and Rapp, D. N. (2014). Retrieval (sometimes) enhances learning: Performance pressure reduces the benefits of retrieval practice. *Appl. Cognit. Psychol.* 28, 597–606. doi:10.1002/acp.3032

Kang, S. H. K., and Pashler, H. (2014). Is the benefit of retrieval practice modulated by motivation?. *J. Appl. Res. Mem. Cognition* 3, 183-188. doi:10.1016/j.jarmac.2014.05.006

Karpicke, J. D., and Roediger, H.L. (2008). The critical importance of retrieval for learning. *Science* 319, 966–968. doi:10.1126/science.1152408

Kornell, N., and Bjork, R. A. (2007). The promise and perils of self-regulated study. *Psychon. B. Rev.* 6, 219-224. doi:10.3758/BF03194055

Kornell, N., Bjork, R. A., Garcia, M. A. (2011). Why tests appear to prevent forgetting: A distribution-based bifurcation model. *J. Mem. Lang.* 65, 85–97. doi:10.1016/j.jml.2011.04.002

Legault, L., and Inzlicht, M. (2013). Self-determination, self-regulation and the brain: Autonomy improves performance by enhancing neuroaffective responsiveness to self-regulation failure. *J. Pers. Soc. Psychol.* 105, 123-128. doi:10.1037/a0030426

Lepper, M. R., Corpus, J. H., Iyengar, S. S. (2005). Intrinsic and extrinsic motivational orientations in the classroom: Age differences and academic correlates. *J. Educ. Psychol.* 97, 184-196. doi:10.1037/0022-0663.97.2.184

Murayama, K., & Kitagami, S. (2014). Consolidation power of extrinsic rewards: reward cues enhance long-term memory for irrelevant past events. *J. Exp. Psychol. Gen.* 143, 15-20. doi: 10.1037/a0031992

Murayama, K., Matsumoto, M., Izuma, K., and Matsumoto, K. (2010). Neural basis of the undermining effect of extrinsic reward on intrinsic motivation. *P. Natl. Acad. Sci. USA* 107, 20911-20916. doi:10.1073/pnas.1013305107

Nielson, K. A., & Bryant, T. (2005). The effects of non-contingent extrinsic and intrinsic rewards on memory consolidation. *Neurobiol. Learn. Mem.* 84, 42-48. doi:10.1016/j.nlm.2005.03.004

Roediger, H. L., and Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends Cogn. Sci.* 15, 20–27. doi:10.1016/j.tics.2010.09.003

Roediger, H. L. and Karpicke, J. D. (2006a). The power of testing memory: Basic research and implications for educational practice. *Perspect. Psychol. Sci.* 1, 181–210. doi:10.1111/j.1745-6916.2006.00012.x

Roediger, H. L., and Karpicke, J. D. (2006b). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychol. Sci.* 17, 249–255. doi:10.1111/j.1467-9280.2006.01693.x

Ryan, R. M., and Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* 55, 68–78. doi:10.1037/0003-066X.55.1.68

Simons, J., Dewitte, S., and Lens, W. (2004). The role of different types of instrumentality in motivation, study strategies, and performance: Know why you learn, so you'll know what you learn!. *Brit. J. Educ. Psychol.* 74, 343-360. doi:10.1348/0007099041552314

Vansteenkiste, M., Simons, J., Lens, W., Sheldon, K. M., and Deci, E. L. (2004). Motivating learning, performance, and persistence: The synergistic effects of intrinsic goal contents and autonomy-supportive contexts. *J. Pers. Soc. Psychol.* 87, 246–260. doi:10.1037/0022-3514.87.2.246

Figure Legends

Figure 1. Results of the Experiment. (A) Probability of correct recall and commission errors in the immediate memory test as a function of extrinsic motivation (low, high). **(B)** Probability of correct recall in the delayed long-term memory test as a function of item type (tested, restudied, control) and

extrinsic motivation (low, high). **(C)** Conditional probability of correct recall in the delayed long-term memory test given successful recall in the immediate memory test as a function of extrinsic motivation (low, high). Error bars represent standard errors of the means.

Figure 1.TIF