Autonomy and Task Performance: Explaining the Impact of Grades on Intrinsic Motivation

Caroline Pulfrey
University of Lausanne

Céline Darnon
Clermont University and the French University Institute

Fabrizio Butera
University of Lausanne

The use of grades to motivate constitutes an unresolved theoretical controversy. In 2 experiments carried out with different age groups and academic tracks, a standard-grade condition was compared with a condition in which differential scoring engendered higher grades and with a no-grade condition. The relative power of task performance and task autonomy to explain differences in subsequent intrinsic motivation (self-report task interest and continuing motivation for the task) was assessed. Results show that, compared with the standard-grade condition, both high-grade and no-grade conditions enhanced the 2 forms of subsequent intrinsic motivation. However, although task performance explained higher levels of task interest in the high-grade condition, it failed to explain higher levels of continuing motivation for the task. Task autonomy, conversely, explained the higher levels of both task interest and continuing motivation for the task experienced in the nongraded condition. Results are discussed in the light of an integrative model that differentiates the mediational role of task performance and autonomy, 2 traditional major explanations of the link between grades and intrinsic motivation.

Keywords: grades, intrinsic motivation, self-determination, autonomy, task performance

Should educators regularly use classroom grades to reward achievement? Since the surge of interest in the impact of grading on motivation that manifested itself in the 1970s and 1980s (Butler, 1987, 1988; Butler & Nisan, 1986; Harter, 1978; Maehr & Stallings, 1972), little research has continued exploring this topic. Yet classroom grading is as prevalent if not more so nowadays (Wolf, 2007). Furthermore, intrinsic motivation, a motivation that drives engagement in activities via the inherent satisfaction derived from effective, freely chosen action (Deci, Koestner, & Ryan, 1999a), has never been as necessary for life success as today, when competition for higher education and top jobs is stiffer than ever ("An Unprecedented Admissions Year," 2011; Berry, 2011), and research has shown that productive or deliberate practice, up to 5 hr a day, outweighs natural ability as a robust predictor of outstanding achievement (Ericsson, Krampe, & Tesch-Römer, 1993).

This question also has theoretical significance, pertaining to an ongoing theoretical debate, between proponents of self-determination theory (SDT; e.g., Deci et al., 1999a) and those of general interest theory (GIT; e.g., Eisenberger, Pierce, & Cameron, 1999). The heart of the debate centers on whether the receipt of performance-contingent rewards for initially interesting tasks is more likely to increase subsequent intrinsic motivation, owing to the satisfaction of the need for competence that they provide (Cameron, Banko, & Pierce, 2001; Cameron, Pierce, Banko, & Gear, 2005; Eisenberger & Aselage, 2009), or to decrease it, owing to the lack of satisfaction of the need for autonomy, a result of the external pressure to obtain the reward (Deci et al., 1999a; Ryan & Weinstein, 2009).

Inspired by this debate, the aim of this research was to present an integrative model that tests the relative mediational role of the degree of autonomy experienced in-task and reward-determining task performance in the relationship between grading and intrinsic motivation.

The Grade: A Potential Reward?

The classroom grade constitutes a good example of an ecologically valid performance-contingent reward for multiple reasons. First, the grade reward is a direct linear function of the quality of task performance (see Deci, Koestner, & Ryan, 1999b). Second, grades provide clear diagnostic performance-related information (Darnon, Dompnier, Delmas, Pulfrey, & Butera, 2009; Elliot & Moller, 2003), and thus are potentially able to satisfy competence

needs. Third, grades hold strong socially symbolic value as they represent a competence-related cue, a significant predictor in the reward–motivation relationship (Harackiewicz, Manderlink, & Sansone, 1984). As such, they create value in ensuing competence-related information, the core of the reward's motivating role (Eisenberger et al., 1999). However, grades can also be perceived as controlling by students (Ryan & Weinstein, 2009), thwarting autonomy needs. As such, they accurately reflect the competing tendencies characteristic of performance-contingent rewards both to control but at the same time provide affirmation of competence (Deci et al., 1999a) and therefore provide a suitable vehicle with which to examine the performance-contingent reward and motivation question.

Performance-Contingent Rewards: A Source of Motivation?

According to SDT (Deci et al., 1999a), receiving rewards for high performance quality generates conflicting effects, simultaneously satisfying the need for competence necessary for intrinsic motivation and frustrating the need for autonomy, also necessary for intrinsic motivation. According to GIT (Eisenberger et al., 1999), rewards supplied for meeting a performance standard communicate a sense of achievement, satisfying the individual's need for competence and inducing a favorable attitude toward the task. As these types of rewards are associated with success in task performance, the competence-related information provided will "override the controlling aspect of reward" (Cameron et al., 2001, p. 26).

The key difference between these two theories centers on the relative importance attached to successful task performance in the determination of subsequent intrinsic motivation. Yet, there is common ground in the meta-analytical findings. Deci et al. (1999a; Deci, Koestner, & Ryan, 2001), comparing a no-feedback, no-reward control condition with ones in which various categories of performance-contingent rewards were administered, found no significant effects of performance-contingent rewards on self-reported task interest, the degree to which participants found the task they have accomplished interesting and enjoyable (Deci et al., 1999a). On the contrary, comparing the same conditions, they did find that performance-contingent rewards, even the receipt of maximum rewards, undermined free-choice time on task (FCTT), a second measure of intrinsic motivation that records and compares time spent on the similar tasks with that just completed in a free-choice period during which the participant is ostensibly waiting for the experimenter to return from some errand (Deci et al., 1999a).

Eisenberger et al. (1999), for their part, found that performance-contingent rewards when compared with a no-feedback, no-reward control condition, had a positive effect on task interest and no effect on FCTT, a result similar to that of Cameron et al. (2005). Cameron et al. (2001) and Cameron and Pierce (2002) again found a positive effect for task interest but a negative effect for FCTT. Both sets of results differentiate the effects of rewards on task interest versus FCTT, with performance-contingent rewards, when compared with a no-feedback, no-reward control condition, having a relatively more positive effect on task interest than on FCTT.

Two Distinct Types of Intrinsic Motivation?

Although the two measures are often aggregated under the umbrella term of *intrinsic motivation* (Cameron et al., 2001; Deci et al., 1999a; Eisenberger et al., 1999), the very fact that many studies include these two measures and that rewards seem to impact them differentially implies that two conceptually and empirically distinct but related types of subsequent intrinsic motivation exist, one more focused on reactions to the task accomplished (puzzle enjoyment, Harackiewicz, 1979; task enjoyment, Harackiewicz, Abrahams, & Wageman, 1987; intrinsic interest, Eisenberger & Aselage, 2009; and free-choice task interest, Cameron et al., 2005) and the other more focused on task persistence once the initial task is ended (continuing motivation for the task, Maehr & Stallings, 1972; continuing task involvement, Butler, 1987; sustained involvement, Patrick, Skinner, & Connell, 1993; and long-term motivation, Wicker, Brown, & Paredes, 1990). The latter, namely continuing motivation, has been operationalized either as FCTT—more common in laboratory studies (see Harackiewicz et al., 1984; W. D. Pierce, Cameron, Banko, & So, 2003, for example)—or, in the case of classroom-based studies, variously as extra tasks requested (Butler, 1987, 1988; Butler & Nisan, 1986), or other measures of intention, such as Maehr and Stalling's (1972) scale of increasing intentional commitment. The common point of both types of measure is their posttask, continuing orientation that denotes voluntary ongoing engagement in the task or a comparable task.

Other literature indicates that these two components of intrinsic motivation might not follow the same pattern of relations with predictors. Elliot et al. (2000) pointed out that self-report and behavioral persistence measures of intrinsic motivation are commonly correlated when feedback is confirming, but often uncorrelated when nonconfirming feedback is provided. Wicker, Brown, Wiehe, and Shim (1990) and Wicker, Brown, and Paredes (1990) note the lack of consistently significant correlations between task interest and FCTT and find differences in their relations with other motivational indices. Likewise, in the literature on grades and intrinsic motivation, differences emerge. Butler's (1988) research found that, in comparing high and low performance, the high grades associated with successful task performance generated higher levels of task interest as long as future tasks of the same sort were deemed to be impending. However, once no more grades were anticipated, there was a significant drop in task interest compared with baseline and previous measures of it. Furthermore, a separate measure of continuing motivation for the task, namely, the number of tasks of the same sort requested at the end of the experiment by the student, was also shown to have declined in the graded conditions compared with the same baseline measure taken before grading was introduced. In addition to this, although task interest remained as high for high-achieving performance in the graded conditions as in the nongraded condition as long as further graded feedback was expected, both postexperiment task interest as well as the number of extra tasks requested were significantly lower in the graded conditions compared with nongraded conditions. These results first reinforce the hypothesis that there might indeed be a distinction between a more proximal, reactive motivation and a more distal ongoing one. Thus, the first contribution of the present research is the presentation of an integrative model that treats posttask intrinsic motivation as two distinct, though

related, constructs, namely, self-report task interest and enjoyment (task interest) and continuing motivation for the task.

Second, they reflect the findings of the meta-analyses reported above: When task performance is sufficiently high to generate a rewarding grade, it has a positive effect on reactions to the task. However, for ongoing forms of motivation, the positive effect of graded high-task performance is no longer apparent. This latter finding reflects the results of Lepper, Greene, and Nisbett's (1973) classic study that revealed that rewarding preschoolers with "Good Player Awards," arguably the preschool equivalent of a good grade, had an undermining effect on subsequent motivation measured as time spent on the same task 7–17 days later when compared with a control group who received no reward. Consequently, we ask the question: Could two different intermediary driving mechanisms be at work?

The Different Mediational Roles of Task Performance and Task Autonomy

In the case of the performance-contingent reward, the intermediate driver of subsequent intrinsic motivation is the task performance attained, be it experimentally manipulated or measured on a scale. From a GIT perspective, when the reward is explicitly tied to a successful performance standard, high task performance predicts task interest and enjoyment because of the achievement-related satisfaction (Eisenberger et al., 1999) that ensues from public recognition of successful performance (see Elliot et al., 2000). From an SDT perspective, the satisfaction of the need for competence (Deci & Moller, 2005) is equivalent to GIT's achievement-related satisfaction. However, conditioning theory (Skinner, 1938) established that the removal of a regular reward led to response extinction. Likewise, in the case of performance-contingent rewards, it could be that the reward reinforcement does not have a significant impact on intentions to continue the activity when no further reinforcement is expected.

In the case of the no-grade/no-reward conditions, the intermediate driver of subsequent intrinsic motivation is the experience of task autonomy in the accomplishment of the task. DeCharms (1968) argued that when *induced* to take part in ego-involving activities, failure to achieve goals threatens self-esteem, and, therefore, ego-involving inductions are likely to be perceived as autonomy reducing. Ryan (1982), too, posited that the more we perceive communications from others as pressure to attain a specific outcome, the less likely we are to experience task autonomy. Thus, task autonomy is likely to be negatively affected by the anticipation of performance-contingent rewards, relative to no-reward conditions.

When it comes to the impact of task autonomy on various forms of subsequent intrinsic motivation, Self-determination proponents argue that "mediation of the undermining effect by the diminution of perceived self-determination" (Deci et al., 1999b, p. 694) is central to the reward–intrinsic motivation link (Deci et al., 2001), as satisfying an individual's need for autonomy is essential to generate intrinsic motivation (Ryan & Deci, 2006). Earlier effectance motivation theory conceptualized motivated organisms as "self-governing entities" constantly seeking further competence and autonomy via repeated bouts of engagement in a task (White, 1959, p. 324)—in other words, continuing motivation for the task.

Thus, the second contribution of the integrative model elaborated in the present research is the testing, within the ecologically valid context of a school context, of the mediational role of task performance and of perceived task autonomy on the relationship between grading and both self-report task interest and continuing motivation for the task.

Hypotheses, Methodological Issues, and Overview

The proposed model, presented in Figure 1, contends that both perceptions of autonomy experienced while accomplishing a task and reward-contingent task performance should mediate the relationship between the grading context of an exercise and self-report task interest (task interest). However, task autonomy and not task performance should mediate the relationship between the grading context of an exercise and continuing motivation for the task. In order to test these two general hypotheses within a classroom context, we used a set of three conditions producing two contrasts with a standard grading condition as the reference condition. In our first contrast, the relative *reward contrast*, this standard-grade condition, characterized by the relatively stiff grading scheme typical of school grading, was pitted against another graded condition, in which a more lenient scoring system would generate a range of generally higher scores. This first contrast enabled us to examine experimentally to what extent higher levels of task performance represented by an experimentally induced higher grade could explain higher levels of subsequent motivation (task interest and continuing motivation for the task). The second contrast, the relative *autonomy contrast*, in which the standard-grade condition was pitted against a nongraded condition, permitted us to examine to what extent higher levels of task autonomy could explain higher levels of subsequent motivation.

It should be noted that in task performance, the meaning of the performance attained is determined by external social agents (e.g., the teachers), and thus the manipulation of task performance directly impacts the reward obtained (the grade). In the majority of past experiments testing the relation of rewards on intrinsic motivation, whether task performance was manipulated (Harackiewicz et al., 1984, Study 3), or was given a meaning by the relative norms of performance (Harackiewicz et al., 1984, Studies 1 and 2), or was measured relative to a performance standard that guaranteed success (Harackiewicz, 1979), it always generated a reward in the reward conditions. Our research design varies from traditional experiments in two ways. First, with the baseline "standard grade" condition, we have the opportunity to experimentally compare the effects of lower versus higher task performance without directly manipulating the scores ourselves. This contrast made it possible to cover the whole range of grades, even negative ones generating an experience widespread in the real world and likely because only partial or no rewards are received, to negatively impact intrinsic motivation (Deci et al., 1999a, 1999b, 2001; Pritchard, Campbell, & Campbell, 1977). Second, we measured task performance as a continuous variable, a practice recommended by Deci et al. (1999b, p. 695) as more "ecologically valid," as the reward people receive is a direct linear function of the quality of their performance.
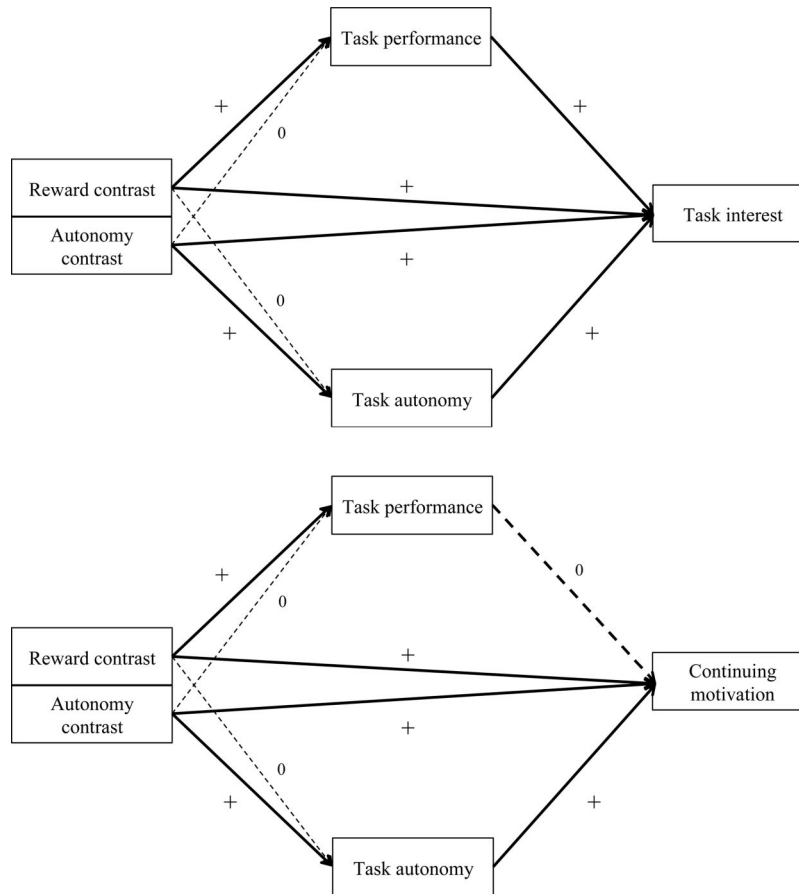
Figure 1.   The theoretical model: Predicted direct and indirect effects of the reward contrast (high grade vs. standard grade) and the autonomy contrast (no grade vs. standard grade) on task interest and continuing motivation for the task. Plus signs indicate significant positive relations ($p < .05$), and zeroes indicate nonsignificant relations ($p > .05$). Solid lines = significant relationships; larger dashed lines = nonsignificant relationships; smaller dashed lines = nonpredicted by the theoretical model.

Thus, our first hypothesis contends that a condition in which no grades are attributed will generate higher levels of *task interest* compared with a standard-grading condition, and this difference will be mediated by the higher levels of task autonomy experienced in the no-grade condition. In addition, a condition in which higher (i.e., more rewarding) grades are generated by means of a more generous grading scheme will also generate higher levels of *task interest* compared with a standard-grading condition, a difference mediated by higher task performance, which procures the more rewarding grade.

The second hypothesis contends that a condition in which no grades are attributed will generate higher levels of *continuing motivation for the task* compared with a standard-grading condition and that this difference will be mediated by the higher levels of perceived task autonomy experienced in the no-grade condition. On the contrary, the higher levels of *continuing motivation for the task* likely to be associated with a condition in which higher than standard grades are generated relative to a standard-grading condition (see Maehr & Stallings, 1972) will not be the result of the higher performance levels generated by the higher grade scheme condition. No mediation effect will appear in this case.

In Experiment 1, we tested our hypotheses among pupils of the same academic track in the 3 years constituting middle school in Switzerland (Grades 7 through 9), and Experiment 2 was designed to assess the generality of these hypotheses with pupils of varying academic abilities (three academic tracks spanning the entire academic range).

Experiment 1

Method

Participants.   Eighty-nine pupils in seventh to ninth grade at a public secondary school in a medium-size French-speaking town in Switzerland participated in the experiment. They were all enrolled in the lowest, namely, the least demanding academically, of the three tracks of this school system. The sample consisted of 53 boys and 36 girls, with a mean age of 15.01 ($SD = 1.18$). The school principal, the teachers, and the parents had given their consent beforehand.

Procedure.

Training session and preparation.   The experiment was carried out as part of a regular class and was presented to the pupils

with the collaboration of the class teacher as a French vocabulary exercise. The classes who participated were brought together in a single group and randomly allocated to one of the three experimental conditions. Three experimenters, trained to provide exactly the same information in each case, ran the experiment simultaneously in three adjacent rooms. The pupils were informed that the activity was organized jointly by the University and their school. Furthermore, they were told that by filling in the questionnaires, they would contribute to a research project for which individual data would be used anonymously. Pupils were given the option not to fill in the questionnaires, and no one refused. The first part of the exercise, the training session, was identical for all three conditions. Each pupil received four envelopes, each containing 10 anagrams, one envelope of three letter anagrams, one of four, one of five, and one of six, and, in addition, an answer sheet. The envelopes had the respective length of the anagrams they contained clearly marked on them. Instructions on how to solve anagrams were provided and an example given on the blackboard. The pupils were then instructed to try out five anagrams, which they could take from any of the envelopes, and were told that they would have 2 min to solve each anagram. This first phase was presented as a means for the pupils to familiarize themselves with the exercise and to try out anagrams of varying difficulty. As further preparation, the pupils were required to fill out on the answer sheet the number of each anagram chosen and the word they had found during the 2 min.

Once the training session was over, the main session was introduced. In each condition, pupils had eight anagrams to solve within the same time limit of 2 min per anagram. Pupils in the no-grade condition were told: "Doing anagrams helps us develop reading and spelling skills. So try to solve as many anagrams as possible and have fun." In the two grade conditions, pupils were told: "Doing anagrams helps us develop reading and spelling skills. That's why, with the collaboration of your teacher, you will receive a grade for the exercise that you're about to do, which will count towards your GPA in French."

The higher and lower grading scheme. Following the example of Harter (1978), in the high-grade condition, students were informed that the promised grade was to be uniquely related to the number of anagrams solved correctly. Each correctly solved anagram was worth 1 point. Eight correctly solved anagrams would earn the pupil a top grade of 100%, and four anagrams correctly solved would earn the pupil a passing grade of 50%. In the standard-grade condition, students were informed that the promised grade was to be a function of the number of anagrams solved correctly multiplied by their respective lengths. The maximum grade of 100% would be attributed for a top score of 48 points (eight anagrams of six letters), and the passing grade of 50% would be obtained with a raw score of 24 points. This more rigorous marking scheme corresponded to the "difficult" evaluation condition used by Maehr and Stallings (1972) and was set up as such to ensure that the reinforcement of the grade would be more negative than in the higher grade condition for all abilities. As in previous performance-contingent reward research, the externally determined meaning of task performance and the resulting representation of the degree of success is what is manipulated, not actual task ability. Although in the two graded conditions pupils were clearly informed, both orally and in writing on their activity pamphlet and

on the blackboard, of the grading scheme, in the no-grade condition no mention of scoring was made.

Measuring task autonomy. At the beginning of the main anagram-solving session, following the announcement of the experimental condition, the pupils filled out the Perceived Locus of Causality (PLOC) questionnaire (Ryan & Connell, 1989), translated into French and adapted to task level. The PLOC questionnaire, constructed on the theoretical basis of personal causation (DeCharms, 1968), measures the participant's self-reported reasons for engaging in a certain behavior and consequently represents with its simplex structure the underlying dimension of task autonomy as experienced by the participant, namely, the degree to which the task is perceived by the participant as supporting their need for autonomous action (Ryan & Deci, 2000). Thus, we can distinguish between internal versus external causality as perceived by the actor, a distinction deemed "crucial for studies of intrinsic motivation versus extrinsic motivation and of task autonomy more generally" (Ryan & Connell, 1989, p. 749).

Main session. During the main session, the pupils tried to solve eight anagrams, each time selecting an anagram out of one of the marked envelopes, noting the length of the anagram on the answer sheet and then trying to solve the anagram. It is worth noting that the pupils were free to choose the anagram length they preferred to attempt, thus providing a certain amount of *within-task* autonomy and the opportunity to try out different strategies to succeed. This reflects the reality of the majority of test-taking situations, in which choices between optional questions and decisions concerning how to tackle the exercise can significantly influence final task scores (Bruch, 1981; A. D. Cohen, 1998; Zhang, Liu, Zhao, & Xie, 2011). At the end of the 2-min period allocated for each anagram, the pupil noted his or her response on the answer sheet and chose the next anagram to solve. Note that with anagram tasks, knowledge of results is immediately evident to the task-taker, as he or she is aware at the end of each 2-min period of whether the word has been found or not. At the end of the activity, a questionnaire measuring task interest and continuing motivation for the task was filled in. Once all the classes had finished the activity, the pupils were debriefed, and those in the grade conditions were told that they could choose whether to keep the grade for their grade-point average (GPA) or not.

Measures.

Task autonomy (RAI). The measure of task autonomy experienced, or the reported autonomy index (RAI; Grolnick & Ryan, 1987), was derived from the four scales of the PLOC questionnaire (Ryan & Connell, 1989), each consisting of five items, translated into French using two-way blind translation with native speakers and adapted to in-class task level as the original items focused on homework. This French version has also been used by Pulfrey, Buchs, and Butera (2011). It included extrinsic motivation with external regulation (e.g., "I'm going to do the exercise mainly because I have to"; $\alpha = .80$), extrinsic motivation with introjected regulation (e.g., "I'm going to do the exercise mainly because I want the teacher to think I'm a hard-working pupil"; $\alpha = .82$), extrinsic motivation with identified regulation (e.g., "I'm going to do the exercise mainly because I think it will be useful for me"; $\alpha = .88$), and intrinsic motivation (e.g., "I'm going to do the exercise mainly because it's interesting"; $\alpha = .94$), and responses were to be given on a scale ranging from 1 (*not at all true for me*)

to 7 (*very true for me*). These four scales were weighted according to their degree of perceived task autonomy: external regulation (-2), introjected regulation (-1), identified regulation (1), intrinsic motivation (2). The final autonomy score is a product of the addition of the weighted scores and constitutes a continuous variable reflecting less to more task autonomy (Grolnick & Ryan, 1987; Patrick et al., 1993).[1] Descriptive statistics for Study 1 are presented in Table 1. Perceptions of task autonomy were measured in the main session after the experimental induction had been announced and immediately prior to the students starting the main task in order to place the mediating variable in the correct chronological order and to tap into pupils' perceptions of the relative autonomy they were experiencing as they started the task.

Task performance. Actual task performance was measured as a function of the number of anagrams solved correctly multiplied by their respective lengths (number of anagrams completed $\times$ number of letters in each correctly solved anagram), thus tapping into performance quality and quantity. Although task performance differentially impacted the reward in the two grade conditions, the very high correlation between task performance and the raw number of anagrams successfully completed ($r = .87$) indicated that task performance was a meaningful measure of perceived achievement in both standard and high-grading schemes.

Self-report task interest (task interest). Five items translated from the Patterns of Adaptive Learning Study (PALS) scale of intrinsic motivation (Midgley, Kaplan, & Middleton, 2001) using the same procedure as before and adapted from class to task level (e.g., "I enjoyed doing the exercise"; "I found the exercise interesting"; $\alpha = .92$) were used to measure task interest. As with all the motivation scales, pupils responded on a scale ranging from 1 (*not at all true for me*) to 7 (*very true for me*). Note that this measure of intrinsic motivation is indeed different from the Intrinsic Motivation subscale of the RAI index from a conceptual point of view—as task autonomy measures the reason why someone is engaging in an activity, and task interest measures their reactions post hoc to the activity.[2] As the raw mean of task interest showed moderate left skew of -0.55 ($SE = 0.22$), the variable was first reflected, then the square root was calculated and the resulting variable was rereflected to reduce skew to acceptable proportions (Osborne, 2002): -0.04 ($SE = 0.22$).

Continuing motivation for the task. The continuing motivation for the task measure aggregated two constructs. On the one hand, desire to do more at a later date was assessed by two items translated from the scale of continuing motivation for the task used by Maehr and Stallings (1972): Pupils indicated (on a scale ranging from 1 [*totally disagree*] to 7 [*totally agree*]) their desire to do more exercises like this one in their free time (e.g., "Would you like to do more exercises like this in your free time?") and their willingness to work on such exercises at home (e.g., "Would you like to take some anagrams to do at home?"). Answer scales for these two questions ranged from 1 (*not at all*) to 7 (*very much*). On the other hand, continuing commitment was assessed by an open question asking the extent of free-time commitment ("If you'd like to do more of these anagrams in your free time, how many anagrams would you like to take home?"; see Butler's, 1987, measure of extra tasks requested). Items were translated using the same procedure as before. As results for this variable covered a wide range ($0–50$, $M = 6.1$, $SD = 7.18$), this last question was recoded into a new variable in which the choice of zero anagrams

was coded as 1, 1–5 anagrams as 2, 6–10 anagrams coded as 3, and over 10 anagrams as 4. Because the answer scale of this variable differed from those of the other two, standardized scores of the three variables were used to calculate the dimension mean ($\alpha = 0.82$).[3]

A Jarque and Bera (1987) Lagrange multiplier (LM) omnibus test of normality carried out on the scales revealed nonsignificant *p* values ranging from .15 to .29. Srivasta's (1984) test of multivariate skew was nonsignificant ($p = .21$) and Mardia's (1985) test of multivariate kurtosis was also nonsignificant ($p = .23$). Descriptive statistic means and correlations are presented in Tables 1 and 2.

Results

Overview of analyses. We carried out regression analyses using SPSS with Hayes and Preacher's (2011) mediate macro, which allows for multiple independent variables and mediators. This made it possible to control for effects of school year and to assess the effects of the two mediators simultaneously.

As noted above, in order to test our two general hypotheses, three conditions were created—standard-grade, higher grade, and no-grade—thus generating two contrasts, the *reward contrast* and

---

[1] An alternative scoring of task autonomy was created using factor scores for the four subdimensions of autonomy. The extrinsic and introjected regulation scores were summed to create a score of controlled motivation, and the identified regulation and intrinsic motivation scores were summed to create a score of autonomous motivation. The controlled motivation score was subtracted from the autonomous motivation score to generate a score of task autonomy. Results using this version of task autonomy closely reflected those of the original autonomy score.

[2] This conceptual distinction is supported by results of a principal component analysis with varimax rotation, which revealed that pretask intrinsically motivated reasons for engaging in the task and posttask task interest items loaded on two separate factors explaining 80.85% of the variance. The rotated component matrix indicated that primary item loadings ranged from .73 to .93. The correlation between components was .70.

[3] This specific measure of continuing motivation for the task was used in preference to the free-choice behavioral measure for five reasons: First, it is conceptually close to White's (1959) fundamental definition of effectance motivation. Second, the ecological validity of continuing motivation for the task, namely, the desire and intention to do more in one's free time, makes this measure particularly diagnostic and appropriate for the classroom setting where the educator's challenge is to encourage engagement not only *in* the classroom but also *outside* the classroom in the pupils' free time. Third, as Elliot et al. (2000—referring to Ryan, Koestner, & Deci,(1991)—argue, the task interest and behavioral measures of persistence have often been found to be uncorrelated in situations in which feedback that is negative or nonconfirming is administered, and, consequently, in such cases self-report measures are to be preferred. Fourth, the FCTT (free-choice time on task) measure, despite being so extensively used in reward research, has been not without its criticisms, ranging from those of Calder and Staw (1975) to those of Wicker, Brown, Wiehe, and Shim (1990) and Wicker, Brown, and Paredes (1990), who also found a lack of correlation between the FCTT and the task interest measure as well as substantially contrasting relations between the two constructs and other measures of motivational and affective constructs. A similar critique has been put forward by Cameron, Pierce, Banko, and Gear (2005), who claim that the FCTT is not a uniform measure of intrinsic motivation that may in fact involve extrinsic as well as intrinsic motivational factors (see Ryan et al., 1991).

Table 1
Experiment 1 (N = 89): Descriptive Statistics and Zero-Order Correlations Among Variables

	No grade			High grade			Standard grade					
Variable	*M*	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range	(2)	(3)	(4)
Task performance (1)	18.91	6.77	3.00–30.00	24.46	4.76	14.00–34.00	18.86	6.59	5.00–29.00	.01	.29**	.21*
Task autonomy (2)	1.84	4.37	−6.30–12.00	1.85	3.97	−6.00–10.10	−0.62	3.39	−7.50–7.80	—	.61***	.51***
Task interest (3)	2.16	0.40	1.51–2.65	2.16	0.37	1.24–2.65	1.67	0.42	1.00–2.65		—	.56***
Continuing motivation for the task (4)	0.11	0.84	−1.24–1.55	2.67	0.72	−1.24–1.90	−0.56	0.72	−1.24–1.02			—

Note. Task performance, task autonomy = uncentered scores.
* *p* < .05. ** *p* < .01. *** *p* < .001.

the *autonomy contrast.* We used dummy coding to code the experimental contrasts: The standard-grade condition was used as the reference condition, which was to be compared, on the one hand, with the more rewarding higher grade condition (the reward contrast), and on the other with the no-grade condition (the autonomy contrast); thus, the standard-grade condition was coded 0 in the two contrasts. In the reward contrast, the higher grade condition, which, we hypothesized, would provide higher levels of task performance and hence a higher, more rewarding grade, was coded 1 and the no-grade condition 0. In the autonomy contrast, the no-grade condition, which, we hypothesized, would engender higher levels of task autonomy owing to the lack of external evaluation, was coded 1 and the higher grade condition 0.

School year was coded using orthogonal contrast coding. In the first contrast, the *seventh-grade contrast*, the seventh graders were coded 2 and the older students (eighth and ninth graders) were each coded −1. In the second contrast, the *eighth-ninth grade contrast*, the seventh graders were coded 0, eighth graders −1, and ninth graders 1. The focus on seventh graders in the seventh-grade contrast was justified by the fact that seventh grade is the first year of academic-level tracking. See Table 2 for means by conditions.

In all, the model included the two sets of two contrasts and interactions between each contrast, resulting in a total of eight independent variables. The two potential mediators, task autonomy and task performance, were also entered in the model. The model was run first with task interest as the dependent variable and then with continuing motivation for the task as the dependent variable (see Figure 2).

Hypothesis 1: Testing the mediation model on task interest. Results of the model with task interest as the dependent variable, presented in Table 3, revealed first in terms of total effects (impact of independent variables without inclusion of mediators – *c* paths; Hayes & Preacher, 2011) that the reward contrast had a significant effect on task interest, with participants in the high-grade condition expressing more task appreciation than those in the standard-grade condition. The autonomy contrast, likewise, had a significant impact on task interest, with participants in the no-grade condition expressing more task appreciation than those in the standard-grade condition.

In terms of the relations between independent variables and mediators (*a* paths), the reward contrast had a significant effect on task performance, with participants in the high-grade condition gaining significantly higher scores than those in the standard-grade condition. The autonomy contrast, however, did not have a generalized significant effect on task performance, with students in the no-grade condition not scoring higher than those in the standard-grade conditions.

The autonomy contrast had, as expected, a significant effect on perceived task autonomy, with participants in the no-grade condition reporting higher levels of perceived task autonomy than those in the standard-grade condition. Unexpectedly, the reward contrast also had a significant effect on task autonomy, with students in the high-grade condition experiencing significantly higher levels of autonomy than those in the standard-grade condition.

When it came to the relations between the two mediators and the dependent variable, task interest (*b* paths), results revealed that both task performance and task autonomy positively predicted task interest.

Indirect effects of the reward and autonomy contrasts on task interest (*ab* paths). Indirect effects and effect sizes are presented in Table 4. We see that the indirect effect (*ab*) of the reward contrast (contrasting the standard-grade condition with the

Table 2
Experiment 1 (N = 89): Means and Standard Deviations by Experimental Condition and School Year

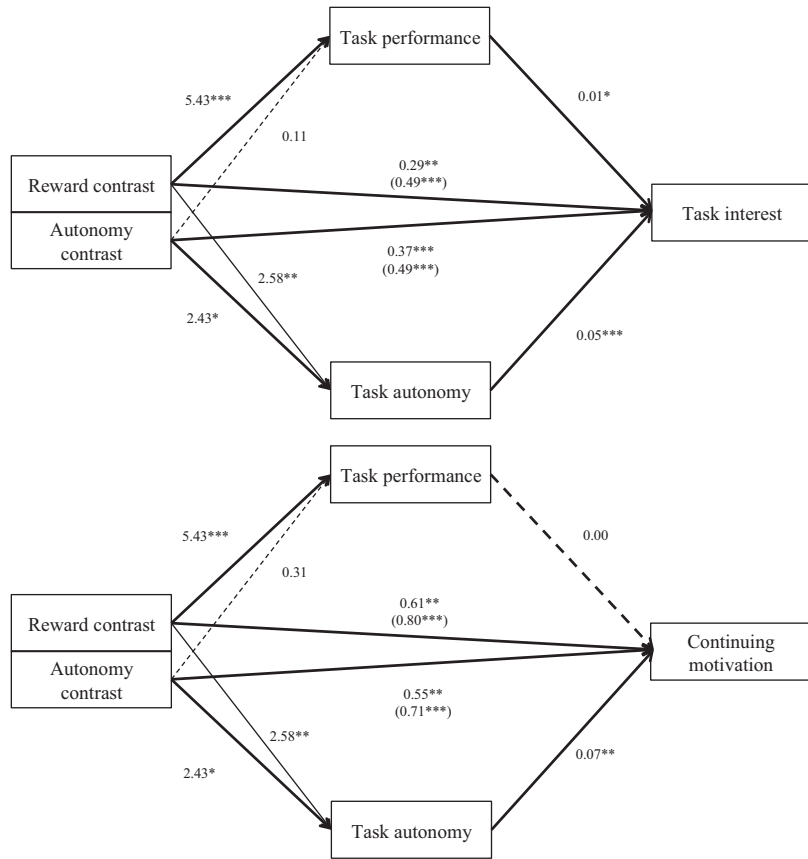|	No grade						High grade						Standard grade					
	7th grade		8th grade		9th grade		7th grade		8th grade		9th grade		7th grade		8th grade		9th grade	
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Task performance	18.46	7.42	20.78	6.24	17.92	6.82	22.33	5.09	25.56	5.70	25.13	3.91	12.09	3.94	21.13	4.26	23.67	4.00
Task autonomy	−0.43	3.15	−0.29	2.65	5.84	3.79	1.80	3.27	3.20	3.76	0.56	5.09	0.25	3.00	−1.74	2.65	−0.70	4.37
Task interest	1.92	0.30	2.07	0.36	2.45	0.35	2.17	0.37	2.23	0.36	2.08	0.42	1.54	0.38	1.70	0.56	1.81	0.32
Continuing motivation for task	−0.48	0.78	0.14	0.65	0.68	0.66	0.23	0.52	0.47	0.88	0.10	0.82	−0.80	0.73	−0.63	0.62	−0.20	0.72

Figure 2. Experiment 1: Measurement model and path diagram of the indirect and direct effects of the reward contrast (high grade vs. standard grade) and the autonomy contrast (no grade vs. standard grade) on task interest and continuing motivation for the task. Solid lines = significant relationships; larger dashed lines = nonsignificant relationships; smaller dashed lines = nonpredicted by the theoretical model. All values represent unstandardized coefficients. $^* p < .05.$ $^{**} p < .01.$ $^{**} p < .001.$

high-grade condition) on task interest via task performance was significant; the direct effect of the reward contrast on task interest (c^1 path) also remained significant. Likewise, the indirect effect of the autonomy contrast (contrasting the standard-grade condition with the no-grade condition) on task interest via task autonomy

was significant, and the direct effect of the autonomy contrast on task interest (c^1 path) remained significant.

Hypothesis 2: Testing the mediation model on continuing motivation for the task. Results of the model with continuing motivation for the task as dependent variable, presented in Table 5,

Table 3
Experiment 1 (N = 89): Model Effects to Task Interest

Path effect	B	SE	F	p	η_p^2
Reward contrast to task interest (c)	0.49	0.10	25.00	.001	.21
Autonomy contrast to task interest (c)	0.49	0.10	26.52	.001	.22
Reward contrast to task performance (a)	5.43	1.51	12.88	.001	.14
Autonomy contrast to task performance (a)	0.11	1.47	0.01	ns	
Reward contrast to task autonomy (a)	2.58	0.97	7.13	.01	.08
Autonomy contrast to task autonomy (a)	2.43	0.94	6.76	.05	.08
Task performance to task interest (b)	0.01	0.01	4.93	.05	.06
Task autonomy to task interest (b)	0.05	0.01	25.10	.001	.25
Reward contrast to task interest (c^1)	0.29	0.09	9.30	.01	.11
Autonomy contrast to task interest (c^1)	0.37	0.09	18.84	.001	.20

Note. Italicized letters represent path coefficients. Italicized letters in parentheses represent model paths (see, e.g., Preacher & Hayes, 2008).

Table 4
Experiment 1 (N = 89): Indirect Effects and Indirect Effect Sizes

Indirect effect	B	SE	95% CI	ab/σ_y	ab/c
Reward contrast − Task performance − Task interest	0.08	0.04	[0.01, 0.16]	0.12	0.16
Autonomy contrast − Task autonomy − Task interest	0.12	0.05	[0.03, 0.22]	0.26	0.25
Reward contrast − Task performance − Continuing motivation	0.01	0.07	[−0.13, 0.17]		
Autonomy contrast − Task autonomy − Continuing motivation	0.16	0.09	[0.03, 0.38]	0.11	0.23

Note. Indirect effect size ab/σ_y represents the partially standardized indirect effects (MacKinnon, 2008) approach in which the impact of a change in the mediator of size a on the dependent variable (Y) is measured in Y standard deviations. Indirect effect size ab/c represents the ratio of the indirect effect to the total effect (Hayes, 2012). CI = confidence interval.

revealed first in terms of total effects that the reward contrast had a significant effect on continuing motivation for the task, with participants in the high-grade condition expressing more desire to do more of the task in their free time than those in the standard-grade condition. The autonomy contrast, likewise, had a significant impact on continuing motivation for the task, with participants in the no-grade condition expressing more desire to do more of the task in their free time than those in the standard-grade condition. For effects of the reward and autonomy contrasts on the mediators: task performance and task autonomy (*a* paths), please refer to results for Hypothesis 1.

When it came to the relations between the two mediators and the dependent variable, continuing motivation for the task, results revealed that task performance did not predict continuing motivation for the task. Perceived task autonomy did, however, predict higher levels of continuing motivation for the task.

Indirect effects and effect sizes are presented in Table 4. The indirect effect of the reward contrast (contrasting the standard-grade condition with the high-grade condition) on continuing motivation for the task via task performance was not significant, whereas the direct effect of the reward contrast on continuing motivation for the task (c^I path) was significant (see Table 5). The indirect effect of the autonomy contrast (contrasting the standard-grade condition with the no-grade condition) on continuing motivation for the task via task autonomy was, however, significant. The direct effect of the autonomy contrast on continuing motivation for the task (c^I path) remained significant.

Discussion

As a whole, these results provide preliminary evidence that whereas both task performance and perceived task autonomy seem to play a role in mediating the relationship between grading context and task interest, perceived task autonomy plays the main mediational role in the relationship between grading context and continuing task motivation. First, although partial mediation of the grading context and task interest relationship by task performance occurred in the high-grade versus standard-grade contrast (the reward contrast), the mediation effect of task performance on the more distal measure of continuing motivation for the task was not significant, with the beta coefficient of the direct effect of task performance on the dependent variable being close to zero. Perceptions of task autonomy, on the contrary, partially mediated the impact of the no-grade versus standard-grade condition on both task interest and continuing motivation for the task.[4]

A limitation of this sample needs to be addressed, however. The sample, being drawn from the lowest academic track in the Swiss secondary system, is likely to be composed almost entirely of pupils who have a history of school failure and difficulty. Considering the potent effects of motivational states and attitudes (Vanlede, Philippot, & Galand, 2006), this might create a bias as pupils officially classified as low achieving might well suffer more from the pressure of a standard-grade condition and conversely be more receptive to self-affirming feedback of a high-grade condition than the average-achieving pupil.

Experiment 2

To ascertain the generalizability of our central hypotheses, a second study was conducted. The study compared the seventh-grade pupils from the lowest academic track who figured in Experiment 1 with seventh-grade pupils from the two higher tracks, thus covering the entire competence range of the Swiss secondary school system in the first year after selection and tracking.

Experiment 2 replicated Experiment 1, except for the addition of measures of task-level self-efficacy, pretask anagram-solving performance, length of anagram chosen in the main task as statistical controls, and a further measure of continuing motivation for the task. Task-level self-efficacy was included as self-efficacy beliefs are argued to significantly affect the motivating impact of outcome expectancies (Bandura, 1995) and also to predict persistence (Lent, Brown, & Larkin, 1987). In addition, as participants in this experiment came from different academic tracks, it was deemed worthwhile including a measure of task-level self-efficacy as a statistical control. Pretask anagram-solving performance and length of anagram chosen in the main task were included to control for varia-

---

[4] Moderation effects by school year, although not part of our main hypotheses, showed two interesting effects. First, although not surprisingly the seventh graders had a lower overall task performance compared with that of the eighth and ninth graders, their scores plummeted in the more challenging standard-grade condition compared with parity of scores in the other two conditions. In addition, they experienced less task autonomy compared with their elders in the no-grade condition. Both of these effects could potentially be outcomes of what Liu, Wang, and Parkins (2005, p. 578) term the "immediate stigmatizing effect" of tracking, in which cross-track comparisons as opposed to within-track comparisons are still salient in individuals' minds when entering the tracking system (i.e., in seventh grade). Further details of these moderating effects and the data sets are available from the authors for interested readers.

Table 5
Experiment 1 (N = 89): Model Effects to Continuing Motivation

Path effect	B	SE	F	p	η_p^2
Reward contrast to continuing motivation (*c*)	0.80	0.19	18.15	.001	.19
Autonomy contrast to continuing motivation (*c*)	0.71	0.18	14.98	.001	.14
Reward contrast to task performance (*a*)	5.43	1.51	12.88	.001	.10
Autonomy contrast to task performance (*a*)	0.11	1.47	0.01	ns	
Reward contrast to task autonomy (*a*)	2.58	0.97	7.13	.01	.06
Autonomy contrast to task autonomy (*a*)	2.43	0.94	6.76	.05	.06
Task performance to continuing motivation (*b*)	0.00	0.01	0.02	ns	
Task autonomy to continuing motivation (*b*)	0.07	0.02	11.29	.01	.13
Reward contrast to continuing motivation (c^1)	0.61	0.20	9.24	.01	.11
Autonomy contrast to continuing motivation (c^1)	0.55	0.18	9.36	.01	.11

Note. Italicized letters represent path coefficients.

tions in task performance and motivational measures resulting from a task-specific pretask level of success as well as the strategy chosen in the choice of anagrams, as this could influence outcome variables independent of pupil ability.

The further measure of continuing motivation for the task consisted of pupils providing their name and address if they wished to receive more anagrams to work on at home. This extra measure of intention and commitment was inspired by Maehr and Stallings' (1972) most extreme measure of continuing motivation for the task in which pupils who had indicated their willingness to participate in another session of task-solving then provided their name and a time when they would be available. The relatively high degree of commitment and intention to pursue the activity predict a significant likelihood of behavioral follow-up (Ajzen, 1991).

Method

Participants and procedure. One hundred twenty-two seventh-grade pupils, enrolled in the three tracks of a secondary school in a medium-size French-speaking town in Switzerland, participated in the experiment. The sample consisted of 70 boys and 52 girls, with a mean age of 13.63 (*SD* = 0.70). The classes that participated were brought together in a single group and randomly allocated to one of the three experimental conditions. Three experimenters, trained to provide exactly the same information in each case, ran the experiment simultaneously in three adjacent rooms. The experimental procedure and the materials were the same as those used in Experiment 1, with the exception that the measure of task-level self-efficacy was taken immediately after the practice session and before the announcement of the experimental conditions; the second measure of continuing motivation for the task, contact details, was added in after the first measure of continuing motivation for the task.

Measures.

Perceived task autonomy (RAI). Alphas for extrinsic motivation with external regulation (α = 0.69), extrinsic motivation with introjected regulation (α = 0.77), extrinsic motivation with identified regulation (α = 0.83), and intrinsic motivation (α = 0.91) were all satisfactory, and once again, these four scales were weighted according to their degree of perceived task autonomy, and the final self-determination score was calculated to create a continuous variable reflecting less to more autonomous or self-determined motivation (Grolnick & Ryan, 1987).

Task-level self-efficacy. The scale of task-level self-efficacy was translated from Midgley et al.'s (2001) Scale of Academic Efficacy using two-way blind translation with native speakers and was adapted to measure task-level self-efficacy (e.g., "I think I will succeed in this exercise"; "I'm able to solve the problems in this exercise"; α = 0.86).

Baseline performance. Students' task performance on the premanipulation introductory exercise in which students attempted to do one three-letter, one four-letter, one five-letter, and one six-letter anagram was calculated. Baseline performance was calculated by summing the number of letters of anagrams successfully completed.

Choice of anagram length. The average length of anagram chosen in the main task was calculated. This measure also served as a manipulation check indicating that the different assessment conditions differentially affected pupil strategy.[5]

Task performance. The same scoring method outlined in Experiment 1 was used.

Task interest. The same PALS scale of intrinsic motivation (Midgley et al., 2001; α = 0.91) was used to measure task interest.[6] The same scaling correction process was used as in Experiment 1 to correct for negative skew.

Continuing motivation for the task 1. As in Experiment 1, our first measure of continuing motivation for the task was as-

---

[5] Results of an analysis of variance indicated a main effect of assessment condition on length of anagram chosen, $F(2, 122) = 70.24$, $p = .001$. Least significant difference post hoc analyses revealed that pupils chose easier anagrams in the high-grade condition (*M* = 3.63, *SD* = 0.58) compared with the no-grade condition (*M* = 4.45, *SD* = 0.65; $p < .001$), and the standard-grade condition (*M* = 5.07, *SD* = 0.42; $p < .001$). Furthermore, they chose easier anagrams in the no-grade condition ($p < .001$) compared with the standard-grade condition ($p < .001$).

[6] As in Experiment 1, a principal component analysis with varimax rotation revealed that the intrinsic motivation component of the PLOC and posttask task interest items loaded on two separate factors explaining 76.52% of the variance. The rotated component matrix indicated that primary item loadings ranged from .62 to .91. The correlation between components was .67.

sessed by aggregating two continuous measures: desire to do more exercises and number of anagrams requested ($\alpha = 0.75$).[7]

A Jarque and Bera (1987) LM omnibus test of normality carried out on the scales revealed p values ranging from .05 to .87. Srivasta's (1984) test of multivariate skew was nonsignificant ($p = .21$) and Mardia's (1985) test of multivariate kurtosis was also nonsignificant ($p = .32$). Means and correlations for all continuous variables are presented in Tables 6 and 7.

Continuing motivation for task 2: Contact information. Pupils provided their name and address in a space provided at the end of the questionnaire if they wished to receive more anagrams to work on at home (Maehr & Stallings, 1972). This dichotomous measure was coded 1 for those who did not provide their name and address and 2 for those who did. In the sample as a whole, 37.2% participants provided their contact information, and 62.8% did not.

Results

Overview of analyses. We once again carried out regression analyses using SPSS with Hayes and Preacher's (2011) mediate macro, used this time to control for effects of academic track while again assessing the effects of the two mediators simultaneously.

We used the same dummy coding to code the experimental contrasts as in Experiment 1, with the standard-grade condition set as the reference condition coded 0 throughout, compared in one contrast (the reward contrast) with the higher, more rewarding grade condition, coded 1, whereas the no-grade condition was coded 0 and compared in the other contrast (the autonomy contrast) with the no-grade condition, coded 1, whereas the high-grade condition was coded 0.

We coded academic track using orthogonal contrast coding. In the first contrast, the low-track contrast, the lowest track in terms of prior academic achievement, the *Option* track,[8] was coded 2, and the two higher tracks, the *General* track and the *Baccalaureat* track, were coded −1. In the second contrast, the middle-high track contrast, the Option track was coded 0, the middle track, the General track was coded 1, and the highest track, the *Baccalaureat* track, was coded −1. The focus on the lowest track in the first contrast was justified by the fact that the Option track has been the object of controversy concerning the lack of professional future and discrimination facing students in this track (20 min online, 2010; Da, 2010). See Table 7 for means by conditions.

In all, the model included the two sets of two contrasts and interactions between each contrast, resulting in a total of eight independent variables. We included three covariates, task-level self-efficacy, baseline anagram-solving performance, and length of anagram chosen, in the model as statistical controls. We also entered the two potential mediators, task autonomy and task performance, in the model. Given the large number of independent variables, the model was trimmed after preliminary analyses on each dependent variable (Kenny, 2011; Nygreen, 1971) with moderating and control predictors that had no significant effect on any of the dependent variables being removed. This left us with a trimmed model containing seven terms: the reward contrast, the autonomy contrast, the two track contrasts, the interaction between the low versus middle-high track contrast and the reward contrast,

the interaction between the middle-high track contrast and the autonomy contrast, and self-efficacy.

Hypothesis 1: Testing the mediation model on task interest. Results of the model with task interest as the dependent variable, presented in Table 8, revealed first in terms of total effects (Hayes & Preacher, 2011) that the reward contrast had a significant effect on task interest, with participants in the high-grade condition expressing more task interest than those in the standard-grade condition. The autonomy contrast, likewise, had a significant impact on task interest, with participants in the no-grade condition expressing more task interest than those in the standard-grade condition.

Testing the relations between independent variables and mediators, the reward contrast had a significant effect on task performance, with participants in the high-grade condition gaining significantly higher scores than those in the standard-grade condition. As in Experiment 1, the autonomy contrast did not have a significant effect on task performance, with students in the no-grade condition not scoring significantly higher than those in the standard-grade conditions. Task self-efficacy also significantly predicted task performance.

The autonomy contrast had, as expected, a significant effect on task autonomy, with participants in the no-grade condition reporting higher levels of perceived task autonomy than those in the standard-grade condition. In this model, the reward contrast exerted a significant effect on task autonomy, with students in the high-grade condition showing higher levels of autonomy than those in the standard-grade condition.

In terms of relations between the two mediators and the dependent variable, task interest (b paths), results revealed that both task performance and task autonomy positively predicted task interest.

Indirect effects of the reward and autonomy contrasts on task interest (ab paths). As indicated in Table 9, the indirect effect of the reward contrast on task interest via task performance was significant. However, the direct effect of the reward contrast on task interest (c^1 path) remained significant. The indirect effect of the autonomy contrast on task interest via task autonomy was also significant, and the direct effect of the autonomy contrast on task interest (c^1 path) was still significant (see also Figure 3).

Hypothesis 2, Part 1: Testing the mediation model on continuing motivation for the task–Measure 1: Desire and intention to do more anagrams in free time. Results of the model with continuing motivation for the task as the dependent variable,

---

[7] As in Experiment 1, a principal component analysis with varimax rotation revealed that task interest and continuing motivation for the task items loaded on two separate factors explaining 73.51% of the variance. The rotated component matrix indicated that primary item loadings ranged mainly from .80 to .92, with just one negatively framed item in the task interest scale "the exercise did not bore me" with a .55 loading on task interest. As the item's cross-loading was negligible, −.03, it was retained in the original scale. The correlation between components was .59.

[8] The *Option* track, the least demanding academically, directs students over a 3-year period toward the job market; the *General* track, moderately demanding academically, directs students over a 3-year period toward ongoing professional training; and the *Baccalaureat* track, the most academically demanding track, directs students over a 3-year period toward the classical high school honors curriculum (Canton de Vaud: site Officiel, 2011).

Table 6
Experiment 2 (N = 122): Descriptive Statistics and Zero-Order Correlations Among Variables

Variable	No grade			High grade			Standard grade			(2)	(3)	(4)	(5)	(6)	(7)
	M	SD	Range	M	SD	Range	M	SD	Range						
Task-level self-efficacy (1)	4.42	0.88	2.00–6.25	4.05	0.88	1.88–6.00	4.71	0.74	3.13–6.25	0.04	0.10	−0.02	0.16	0.42***	0.38***
Task performance (2)	19.87	6.75	3.00–30.00	23.09	5.74	13.00–38.00	17.9	6.73	5.00–32.00	—	0.13	0.42***	0.22*	0.15	−0.10
Task autonomy (3)	2.47	4.31	−8.30–12.60	0.93	3.87	−5.30–10.30	−0.51	2.91	−6.80–7.80		—	0.54***	0.38***	0.04	−0.04
Task interest (4)	2.13	0.38	1.37–2.65	2.09	0.4	1.20–2.65	1.6	0.4	1.00–2.65			—	0.41***	−0.02	−0.24**
Continuing motivation for the task (5)	0.05	0.93	−1.27–1.58	0.17	0.71	−1.18–1.58	−0.31	0.85	−1.27–1.41				—	−0.05	−0.16
Premanipulation baseline performance (6)	2.97	1.2	0.00–4.50	2.35	0.93	0.75–4.50	3.43	1.09	0.75–4.50					—	0.42***
Mean length of anagram chosen (7)	4.45	0.65	3.00–5.75	3.63	0.59	3.00–4.88	5.07	0.42	4.38–5.88						—

* $p < .05.$ ** $p < .01.$ *** $p < .001.$

presented in Table 10, revealed first in terms of total effects that the reward contrast had a significant effect on continuing motivation for the task, with participants in the high-grade condition expressing more continuing motivation for the task than those in the standard-grade condition. The autonomy contrast, likewise, had a significant impact on continuing motivation for the task, with participants in the no-grade condition expressing more continuing motivation for the task than those in the standard-grade condition. Effects of the reward and autonomy contrasts on the mediators— task performance and task autonomy (*a* paths)—are identical to those in Hypothesis 1. When it came to relations between the two mediators and the dependent variable, continuing motivation for the task (*b* paths), results revealed that task performance did not predict higher levels of continuing motivation for the task, but task autonomy did. Task self-efficacy once again positively predicted task performance and, in addition, also positively predicted continuing motivation.

Indirect effects of the reward and autonomy contrasts on continuing motivation for the task. As shown in Table 9, the indirect effect of the autonomy contrast (contrasting the standard-grade condition with the no-grade condition) on continuing motivation for the task via task autonomy was significant; the direct effect of the autonomy contrast on continuing motivation for the task was no longer significant. On the contrary, although the direct effect of the reward contrast on continuing motivation for the task was significant, the indirect effect of the reward contrast via task performance was not.

Hypothesis 2, Part 2: Testing the mediation model on continuing motivation-Measure 2: Provision of contact information. The model created to test the mediation model on the second measure of continuing motivation for the task, namely, whether students had provided their address or not with a view to receiving more anagrams at home, used the SPSS Indirect macro (Preacher & Hayes, 2008), as this accommodates dichotomous outcome variables. The model was first run with the autonomy contrast entered as a key variable and the other retained independent variables entered as covariates. The model was then run a second time, with the reward contrast entered as a key variable and the other retained independent variables, namely, the autonomy contrast, the middle-high track contrast, and task level self-efficacy, entered as covariates.

Results, presented in Table 11, revealed a similar pattern to those of the first measure of continuing motivation. In the first model, the total effect (*c*) of the autonomy contrast on provision of contact information was significant, and it also had a significant impact on task autonomy. Task autonomy positively predicted the provision of contact information, and this generated an indirect effect between the autonomy contrast and Measure 2 of continuing motivation via task autonomy. The direct effect of the autonomy contrast on provision of contact information was no longer significant. Self-efficacy significantly predicted provision of contact information.

In a second model with the reward contrast as the main independent variable, there was an effect of the reward contrast on provision of contact information and on task performance, but no effect of task performance on provision of contact information. As in the first measure of continuing motivation for the task, there was an effect of the reward contrast on task autonomy.

Table 7

Experiment 2 (N = 122): Means and Standard Deviations by Experimental Condition and Academic Track

	No grade						High grade						Standard grade					
	Lowest track		Middle track		Highest track		Lowest track		Middle track		Highest track		Lowest track		Middle track		Highest track	
Variable	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Task performance	18.46	7.42	20.53	5.77	20.33	7.62	22.33	5.09	21.38	4.50	25.53	6.81	12.09	3.94	17.07	5.71	23.00	5.54
Task autonomy	−0.43	3.15	3.75	5.04	3.53	3.10	1.80	3.27	0.48	3.85	0.69	4.43	0.25	3.00	−1.13	2.88	−0.44	2.95
Task interest	1.92	0.30	2.31	0.43	2.09	0.27	2.17	0.37	1.89	0.32	2.23	0.42	1.54	0.38	1.42	0.27	1.82	0.44
Continuing motivation for task	−0.58	0.72	0.48	1.00	0.10	0.80	0.10	0.45	0.28	0.74	0.12	0.86	−0.85	0.66	−0.34	0.72	0.11	0.90

Discussion

Results of Experiment 2 replicate the main results of Experiment 1, across the three academic tracks, while controlling for task-level self-efficacy, pretask baseline performance, and length of anagram chosen in task, hence confirming our hypotheses that, although both task performance and perceived task autonomy intervene to explain why grading context influences task interest, perceived task autonomy alone holds significant explanatory power when it comes to continuing task motivation. Once again, task performance partially mediated the grading context—task interest relationship in the high-grade versus standard-grade comparison—but showed no significant mediational power in the case of continuing motivation for the task. Task autonomy again provided partial mediation of the impact of grading context (no grade vs. standard grade) on task interest and, in addition, fully mediated its impact on continuing motivation for the task. These results indicate the potential of task autonomy to act as intermediary between graded versus nongraded task conditions and not only immediate task reactions but also, more importantly, desire and concrete intention—as revealed by the effect on both continuing motivation for the task measures—to further engage independently in the activity at a later date. The main effect of task-level self-efficacy on measures of continuing motivation for the task but not of task interest is potentially interesting in that it faithfully reflects certain findings that relate self-efficacy to student effort (Schunk & Hanson, 1985), to persistence (Schunk, 1981), to the seeking of challenging goals (Zimmerman, Bandura, & Martinez-Pons, 1992), and to school engagement, including a subdimension entitled "beyond the call" (Caraway, Tucker, Reinke, & Hall, 2003).[9]

General Discussion

In this research, we wished to revisit and shed new light on the debate concerning the relationship between grades and intrinsic motivation (Butler, 1988, Butler, 1989; Ryan & Deci, 1989). We predicted, first, that task performance and the resulting grade would explain the increases in retrospective self-report task interest experienced in a condition characterized by higher grades compared with one characterized by standard grades and that perceived task autonomy would be able to explain increases in task interest in a condition characterized by no grading compared with one characterized by grades, at equal levels of overall achievement. Our second prediction was, however, that out of the two potential mediators, perceived task autonomy alone would have an impact on continuing motivation for the task (desire and commitment to do more of the task).

Results of two studies, one carried out with pupils from only one academic secondary track and the second carried out with all three academic tracks, revealed that in the comparison between the two graded conditions, although the grade obtained partially explained increased task interest in the higher grade condition, it failed to explain the increased levels of continuing motivation for the task. However, in the comparison between a standard-graded condition and a nongraded condition, which elicited equivalent levels of achievement, higher levels of perceived task autonomy in the no-grade condition partially explained not only the greater task interest experienced in the nongraded condition but also the higher levels of continuing motivation for the task in this same condition.

These results are of particular interest for three different reasons. First, they bring to light new nuances of the old but essentially unresolved grading–intrinsic motivation controversy (Butler, 1989; Ryan & Deci, 1989), characterized by competing explanations of lower continuing motivation for the task after receiving grades

---

[9] Moderation effects of academic track, although not part of our main hypotheses, revealed a linear relation between academic level and task performance, with higher tracks performing better than lower. Motivationally speaking, there were also increases in continuing motivation for the task in higher tracks and increased task interest in the top track compared with the middle track, although, interestingly, it was the middle track who expressed the most task interest in the no-grade condition. This latter effect could be related to the still-prevalent assimilation effects (Oakes, 1985; Trautwein, Lüdtke, Marsh, Köller, & Baumert, 2006), in which cross-level comparisons are still salient in students' minds. Such effects could potentially explain why high-achieving students in the top track might still be more likely to welcome standard-graded exercises than their midrange counterparts, with such exercises still engendering associations of the self-validation that graded exercises furnished them with in previous mixed-ability classes. The corresponding pattern of lower continuing task motivation experienced by the lowest track compared with the middle/upper tracks once again bears out the argument of Liu et al. (2005), in which the negative motivational effects of placement in a stigmatized lower ability track (Slavin, 1988) are likely to be particularly salient in the year following the selection and tracking process, as the focus will still be more on imagined cross-track comparisons than the within-track comparisons with others in the track they are now in. Further details of these moderation effects and the data sets are available from the authors.

Table 8
Experiment 2 (N = 122): Model Effects to Task Interest

Path effect	B	SE	F	p	η_p^2
Reward contrast to task interest (c)	0.55	0.08	41.99	.001	.27
Autonomy contrast to task interest (c)	0.54	0.08	42.64	.001	.27
Reward contrast to task performance (a)	6.31	1.37	21.25	.001	.16
Autonomy contrast to task performance (a)	2.47	1.35	3.39	ns	
Reward contrast to task autonomy (a)	2.06	0.86	5.76	.05	.05
Autonomy contrast to task autonomy (a)	3.21	0.84	14.52	.001	.11
Task performance to task interest (b)	0.01	0.01	7.62	.01	.06
Task autonomy to task interest (b)	0.05	.01	34.93	.001	.24
Reward contrast to task interest (c^1)	0.36	0.08	20.70	.001	.16
Autonomy contrast to task interest (c^1)	0.36	0.08	22.09	.001	.17
Task self-efficacy to task performance	1.33	0.65	4.16	.05	.03

Note. Italicized letters represent path coefficients.

versus no grades thanks to either loss of autonomy or ego involve-ment. Although Butler (1987, 1988) claimed that it was because grading threw the focus of attention onto ego-involvement that good results maintained subsequent intrinsic motivation only as long as more feedback was expected, Ryan and Deci (1989) argued that grading, by situating the PLOC outside the individual in the environment, would erode motivation via loss of autonomy. Our results, showing that in comparisons of graded situations charac-terized by higher versus lower grades the grade received explained differences in self-report task interest but not in continuing moti-vation for the task, provide a potential reconciliation of these two apparently opposing views. Indeed, in support of Butler's asser-tions (Butler, 1987, 1988, 1989), grade quality does influence one aspect of subsequent motivation, namely, task interest. Further-more, this could potentially be related to ego enhancement, as the interest and enjoyment expressed in this appreciation could equally well relate to the receipt of the grade as much as to the content of the task itself (see Deci & Ryan, 1999a). Additionally, the fact that task performance and resulting grade quality failed to explain differences in desire and intention to do more combined with the capacity of perceived task autonomy to explain differences in both immediate task reactions and continuing motivation for the task between a graded and nongraded condition provides support for Ryan and Deci's (1989) arguments that, owing to the controlling nature of grading as a practice, a successful performance that earns a high grade is still not capable of predicting continuing motivation

for the task, whereas perceived task autonomy is (Ryan & Deci, 2006).

The second pole of interest of this research addresses the com-plexity of the intrinsic motivation construct. There is a subtle but important distinction between "being moved to act for the fun or challenge entailed rather than because of external prods, pressures or rewards" (Ryan & Deci, 2000, p. 56) and responding post hoc that one enjoyed an activity having just received a good grade that will boost one's GPA. Consequently, as Butler (1987) argued, higher grades can generate increased retrospective task enjoyment and interest via ego enhancement. However, as Ryan and Deci (1989) argued, it is via the reduction of perceived task autonomy that grades generally reduce "intrinsic" motivation, that is, contin-ued engagement in an activity for its own sake. The explicit distinction between these two aspects of subsequent motivation inherent in the results of our two experiments thus conveniently allows a reconciliation of these two viewpoints and constitutes a step toward a synthesis of the literature on grading and motivation, as well as providing additional empirical confirmation to the argument that two distinct forms of intrinsic motivation exist (Cameron et al., 2005; Elliot et al., 2000; Ryan, Koestner, & Deci,1991; Wicker, Brown, & Paredes, 1990; Wicker et al., 1990).

The third point of interest of this research is that it potentially sheds light on the question of how precisely performance-contingent rewards impact subsequent motivation. Grades carry some of the highest "motivational significance" (Deutsch, 1979, p.

Table 9
Experiment 2 (N = 122): Indirect Effects and Indirect Effect Sizes

Indirect effect	B	SE	95% CI	ab/σ_y	ab/c
Reward contrast − Task performance − Task interest	0.09	0.05	[0.01, 0.19]	0.11	0.16
Autonomy contrast − Task autonomy − Task interest	0.15	0.04	[0.07, 0.24]	0.32	0.28
Reward contrast − Task performance − Continuing motivation 1	−0.01	0.09	[−0.19, 0.17]		
Autonomy contrast − Task autonomy − Continuing motivation 1	0.20	0.09	[0.06, 0.39]	0.23	0.46
Reward contrast − Task performance − Continuing motivation 2	−0.02	0.32	[−.72, .57]		
Autonomy contrast − Task autonomy − Continuing motivation 2	0.41	0.47	[0.01, 1.04]	a	

Note. Indirect effect size ab/σ_y represents the partially standardized indirect effects (MacKinnon, 2008) approach in which the impact of a change in the mediator of size a on the dependent variable (Y) is measured in Y standard deviations. Indirect effect size ab/c represents the ratio of the indirect effect to the total effect (Hayes, 2012).
[a] Both direct and indirect effects are calculated using logistic regression. However, as the indirect and total effects have differential scaling, proportional measures of effect size cannot be calculated (Hayes, 2011).
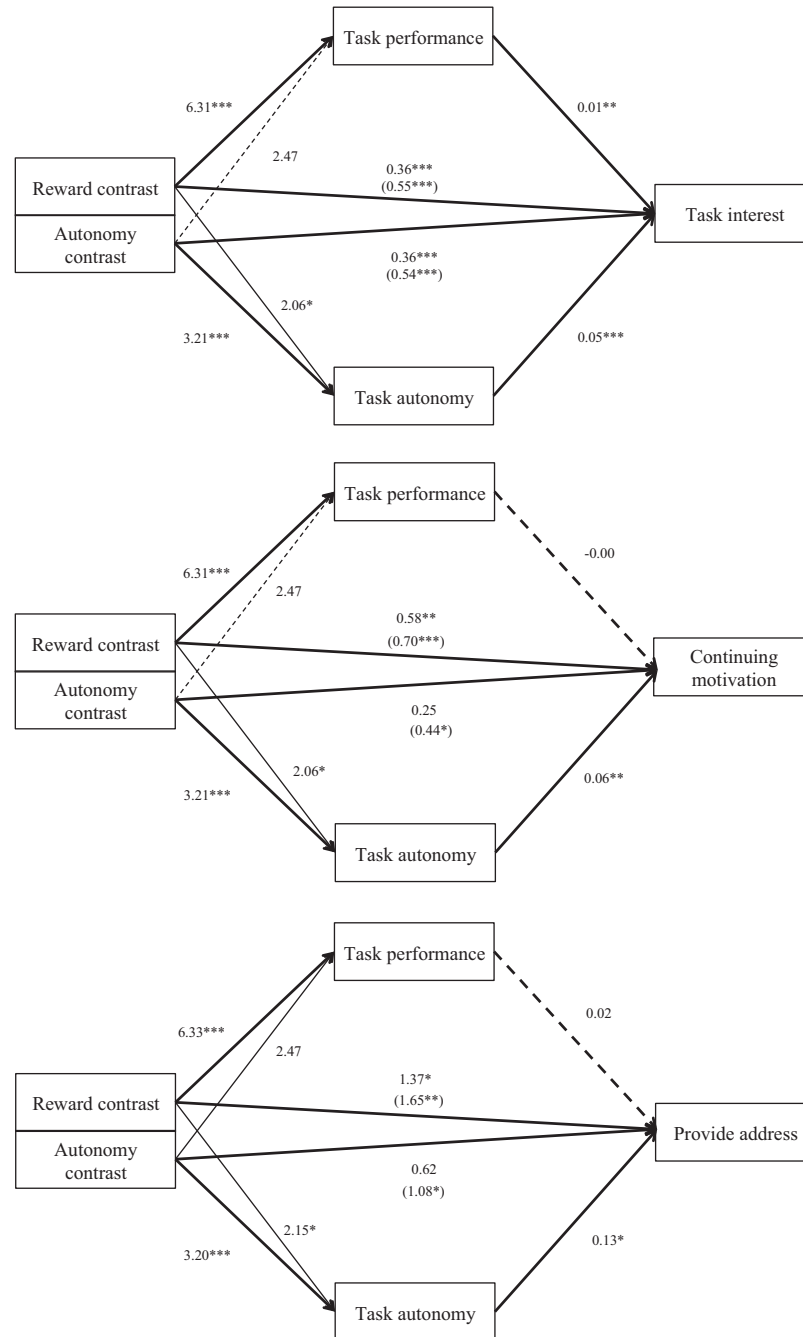
Figure 3. Experiment 2: Measurement model and path diagram of the indirect and direct effects of the reward contrast (high grade vs. standard grade) and the autonomy contrast (no grade vs. standard grade) on task interest, continuing motivation for the task, and providing contact information. Solid lines = significant relationships; dotted lines = nonsignificant relationships; thinner lines = nonpredicted by the theoretical model. All values represent unstandardized coefficients. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

393) of all rewards for those who receive them, with their immediate impact on self-esteem (Butler, 1987, 1989), perception of intelligence (Stapel & Koomen, 2005, Study 2) and performance (Darnon, Harackiewicz, Butera, Mugny, & Quiamzade, 2007), and their more long-term consequences for career advancement (Deutsch, 1979). As such, the results of the two experiments presented in this research, coherent as they are with the main comparative tendencies revealed in meta-analytical findings on performance-contingent rewards (Cameron et al., 2001; Cameron & Pierce 2002; Cameron et al., 2005; Deci et al, 1999a; Eisenberger et al., 1999), have the potential to open up common ground between not only contemporary but also classic behaviorist and

Table 10
Experiment 2 (N = 122): Model Effects to Continuing Motivation 1

Path effect	B	SE	F	p	η_p^2
Reward contrast to continuing motivation (*c*)	0.70	0.18	15.84	.001	.13
Autonomy contrast to continuing motivation (*c*)	0.44	0.17	6.61	.05	.06
Reward contrast to task performance (*a*)	6.31	1.37	21.25	.001	.16
Autonomy contrast to task performance (*a*)	2.47	1.35	3.39	*ns*	
Reward contrast to task autonomy (*a*)	2.06	0.86	5.76	.05	.05
Autonomy contrast to task autonomy (*a*)	3.21	0.84	14.52	.001	.11
Task performance to continuing motivation (*b*)	−0.00	0.01	0.03	*ns*	
Task autonomy to continuing motivation (*b*)	0.06	0.02	11.00	.05	.06
Reward contrast to continuing motivation (*c*[1])	0.58	0.19	9.49	.01	.11
Autonomy contrast to continuing motivation (*c*[1])	0.25	0.18	1.96	*ns*	.09
Task self-efficacy to task performance	1.33	0.65	4.16	.05	.03
Task self-efficacy to continuing motivation	0.25	0.08	9.00	.01	.06

Note. Italicized letters represent path coefficients.

self-determination approaches to the reward-motivation theory. In line with behaviorist theory (Skinner, 1938), increases in immediate posttask reactions can efficiently be explained by the quality of the grade received, a classic case of response reinforcement, but not increases in future intentions once the reinforcement is no longer salient, a case of extinction. Note that in Cameron and Pierce's (2002) most recent meta-analysis, the authors conclude by saying that "when rewards are delivered repeatedly and intrinsic motivation is assessed on more than one occasion, the evidence suggests that the negative effects are fleeting and transitory" (p. 131), thus emphasizing the need for ongoing reinforcement to counter negative effects incurred. In line with SDT, the negative impact on intrinsic motivation caused by the shift of personal causation from internal to external that is characteristic of a reward situation (DeCharms, 1968) and, by extension, of a graded situation compared with a nongraded situation, may, in short-term reactions to the task (or the reward), be offset by pleasure at receiving positive competence-related information, but will still exert its impact on more long-term continuing motivation for the task.

Although we understand people's "frustration with the entire controversy per se" (Lust, 2004, p. 261) when the skeleton of the rewards debate is once again dragged out of the closet, we believe that the results of this research not only shed an alternative light on the debate but also raise a number of intriguing questions, which future research could profitably address. One such question pertains to the explanation of the presence of high levels of continuing motivation for the task in the higher grade condition, a result similar to that of Maehr and Stallings (1972), who, when comparing levels of task difficulty and internal versus external evaluation, found external evaluation on the "easy" task to predict continuing motivation for the task. Continuing motivation for the task was not explained by task performance in either of our experiments, but it could potentially have been explained by motivational states experienced during task accomplishment, for example, lower levels of affective autonomy, namely, the experience of feelings of tension and pressure (Houlfort, Koestner, Joussemet, Nantel-Vivier, & Lekes, 2002), experienced in the higher grade condition compared with the standard-grade condition or, alternatively, a greater context-specific perceived locus of control (Patrick et al., 1993), again experienced in the higher grade condition compared with the lower grade condition. These states have been shown to exert positive effects on subsequent motivation and behavior (Houlfort et al., 2002; Patrick et al., 1993) and could indeed increase the sense of volition associated with autonomous motivation. Further research in this area could usefully clarify whether, when, and why more generous grading schemes might indeed have positive ongoing motivational benefits.

Table 11
Experiment 2 (N = 122): Model Effects to Continuing Motivation 2

Path effect	B	SE	F	Wald	p	η_p^2
Reward contrast to continuing motivation (*c*)	1.65	0.58		7.98	.01	
Autonomy contrast to continuing motivation (*c*)	1.08	0.54		3.96	.05	
Reward contrast to task performance (*a*)	6.33	1.38	21.07		.001	.16
Autonomy contrast to task performance (*a*)	2.47	1.35	3.35		*ns*	
Reward contrast to task autonomy (*a*)	2.15	0.86	6.25		.05	.05
Autonomy contrast to task autonomy (*a*)	3.20	0.84	14.52		.001	.11
Task performance to continuing motivation (*b*)	0.00	0.3		0.30	*ns*	
Task autonomy to continuing motivation (*b*)	0.13	0.06		4.12	.05	
Reward contrast to continuing motivation (*c*[1])	1.37	0.63		2.18	.05	
Autonomy contrast to continuing motivation (*c*[1])	0.62	0.59		1.05	*ns*	
Task self-efficacy to continuing motivation	0.58	0.28		2.07	.05	

Note. Italicized letters represent path coefficients.

Another question concerns the potential quality of free-time free-choice task engagement. Harter (1978) found that desire for challenge was lower in graded than nongraded conditions. Butler (1987, 1988) associated graded activities with a focus on ego enhancement. Ryan et al. (1991) distinguished between free-choice behavior that resulted from ego-involved persistence and that which resulted from genuine intrinsic drives. Could it be that free-time free-choice task engagement resulting from more lenient grading schemes might also serve an ego-reinforcement function as opposed to one of exploration and discovery? This question is of considerable importance in light of the findings that show that practicing an activity leads to ongoing performance improvements only when challenging material is tackled with full concentration and strategic thinking (Ericsson et al., 1993). It is not just how much you do in your spare time that counts, but also how you do it.

A third consideration is that the specificity of the grade as a type of reward/punishment does call for caution in generalizations, and this constitutes a limit to the present analysis. Consequently, replication of this type of paradigm with other types of high-stakes rewards, such as monetary bonuses (Deci & Porac, 1978) or promotion prospects, as well as the more classic low-stakes rewards, such as course credits (Toma & Butera, 2009), token gifts, or food (Deci et al., 1999a), would further contribute worthwhile knowledge to this domain.

A fourth and final concern, emerging from the somewhat small indirect effect sizes, is to what extent the indirect paths explored in this research have the potential to usefully inform classroom practice. Although small, the mediation effects explored are consistent, with main mediation effects holding across age and ability differences and moderation effects accurately reflecting motivational and performance-related fluctuations linked to academic tracking practices. Furthermore, the diminutive nature of these effect sizes might also be related to another limitation of the studies, that is, the use of an anagram-solving task lasting only 50 min. Although related to linguistic skill and a classic task in educational research (Henin, Accorsi, Cho, & Tabor, 2009; Novick & Shermann, 2003), the use of a normal classroom task regularly used for classroom exercise and assessment, for example, reading comprehensions or dictations in French language classes, might engender stronger effects, and as such, further exploration in this area is warranted.

Thus findings, questions raised, and limitations all lead to the conclusion that, at a time when competition for students' free-time attention seems to be at an all-time high, as is indicated by the recent development of an Internet-blocking device that can censor access to Facebook during homework time (B. Cohen, 2011), and when the use of communication technologies by adolescents has been associated with poorer academic performance (T. Pierce & Vaca, 2007), finding ways to manage classroom assessment with a view to unlocking the potential of intrinsic motivation in all its forms is a quest of considerable importance and timeliness.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50,* 179–211. doi:10.1016/0749-5978(91)90020-T

An unprecedented admissions year: Harvard College invites 6.2% of record 34,950 applicants. (2011, March 30). *Harvard Gazette.* Retrieved from http://news.harvard.edu/gazette/story/2011/03/an-unprecedentedadmissions-year/

Bandura, A. (1995). Exercise of personal and collective efficacy in changing societies. In A. Bandura (Ed.) *Self-efficacy in changing societies* (pp. 1–45). New York, NY: Cambridge University Press. doi:10.1017/CBO9780511527692.003

Berry, D. (2011). Ivy-league: The year of single-digit acceptance rates. *College View.* Retrieved from http://www.collegeview.com/admit/?p=1888

Bruch, M. A. (1981). Relationship of test-taking strategies to test anxiety and performance: Toward a task analysis of examination behavior. *Cognitive Therapy and Research, 5,* 41–56. doi:10.1007/BF01172325

Butler, R. (1987). Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest, and performance. *Journal of Educational Psychology, 79,* 474–482. doi:10.1037/0022-0663.79.4.474

Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology, 58,* 1–14. doi:10.1111/j.2044-8279.1988.tb00874.x

Butler, R. (1989). On the psychological meaning of information about competence: A reply to Ryan and Deci's comment on Butler (1987). *Journal of Educational Psychology, 81,* 269–272. doi:10.1037/0022-0663.81.2.269

Butler, R., & Nisan, M. (1986). Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance. *Journal of Educational Psychology, 78,* 210–216. doi:10.1037/0022-0663.78.3.210

Calder, B. J., & Staw, B. M. (1975). Self-perception of intrinsic and extrinsic motivation. *Journal of Personality and Social Psychology, 31,* 599–605. doi:10.1037/h0077100

Cameron, J., Banko, K. M., & Pierce, W. D. (2001). Pervasive negative effects of rewards on intrinsic motivation: The myth continues. *The Behavior Analyst, 24,* 1–44.

Cameron, J., & Pierce, W. D. (2002). *Rewards and intrinsic motivation: Resolving the controversy.* Westport, CT. Bergin and Garvey.

Cameron, J., & Pierce, W. D. (2005). Rewards and motivation in the classroom. *Academic Exchange Quarterly, 9,* 67–71.

Cameron, J., Pierce, W. D., Banko, K. M., & Gear, A. (2005). Achievement-based rewards and intrinsic motivation: A test of cognitive mediators. *Journal of Educational Psychology, 97,* 641–655. doi:10.1037/0022-0663.97.4.641

Canton de Vaud. (2011). *Scolarité obligatoire. Degrés 7 à 9* [Obligatory schooling: Degrees 7 to 9]. Retrieved January, 20, 2012, from http://www.vd.ch/themes/formation/scolarite-obligatoire/

Caraway, K., Tucker, C. M., Reinke, W., & Hall, C. (2003). Self-efficacy, goal orientation and fear of failure as predictors of school engagement in high school students. *Psychology in the Schools, 40,* 417–427. doi:10.1002/pits.10092

Cohen, A. D. (1998). *Strategies in learning and using a second language.* London, England: Longman.

Cohen, B. (2011, May, 9). Talk Talk boss introduces 'homework time' internet controls. *Channel 4 News.* Retrieved from http://blogs.channel4.com/benjamin-cohen-on-technology/talk-talk-boss-introduces-homework-time-internet-controls/2445

Da, M. (2010, March 29). Quels débouchés au sortir de la VSO? *Le Temps.* Retrieved January 30, 2012, from http://www.letemps.ch/Page/Uuid/0e1299b2–3aaa-11df-87f4-7f0bc3738ec3%7C1

Darnon, C., Dompnier, B., Delmas, F., Pulfrey, C., & Butera, F. (2009). Achievement goal promotion at university: Social desirability and social utility of mastery and performance goals. *Journal of Personality and Social Psychology, 96,* 119–134. doi:10.1037/a0012824

Darnon, C., Harackiewicz, J., Butera, F., Mugny, G., & Quiamzade, A. (2007). Performance-approach and performance-avoidance goals: When

uncertainty makes a difference. *Personality and Social Psychology Bulletin, 33,* 813–827. doi:10.1177/0146167207301022

DeCharms, R. (1968). *Personal causation: The internal affective determinants of behavior.* New York, NY: Academic Press.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999a). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125,* 627–668. doi:10.1037/0033-2909.125.6.627

Deci, E. L., Koestner, R., & Ryan, R. M. (1999b). The undermining effect is a reality after all—Extrinsic rewards, task interest, and self-determination: Reply to Eisenberger, Pierce, and Cameron (1999) and Lepper, Henderlong, and Gingras (1999). *Psychological Bulletin, 125,* 692–700. doi:10.1037/0033-2909.125.6.692

Deci, E. L., Koestner, R., & Ryan, R. M. (2001). Extrinsic rewards and intrinsic motivation in education: Reconsidered once again. *Review of Educational Research, 71,* 1–27. doi:10.3102/00346543071001001

Deci, E. L., & Moller, A. C. (2005). The concept of competence: A starting place for understanding intrinsic motivation and self-determined extrinsic motivation. In C. Dweck & E. Elliot (Eds.), *Handbook of competence and motivation* (pp. 529–597). New York, NY: Guilford Press.

Deci, E. L., & Porac, J. F. (1978). Intrinsic motivation and cognitive evaluation theory. In M. Lepper and D. Greene (Eds.), The Hidden Costs of Rewards. Hillsdale, N. J.: Erlbaum.

Deutsch, M. (1979). Education and distributive justice: Some reflections on grading systems. *American Psychologist, 34,* 391–401. doi:10.1037/0003-066X.34.5.391

Eisenberger, R., & Aselage, J. (2009). Incremental effects of reward on experienced performance pressure: Positive outcomes for intrinsic interest and creativity. *Journal of Organizational Behavior, 30,* 95–117. doi:10.1002/job.543

Eisenberger, R., Pierce, W. D., & Cameron, J. (1999). Effects of reward on intrinsic motivation—Negative, neutral, and positive: Comment on Deci, Koestner, and Ryan (1999). *Psychological Bulletin, 125,* 677–691. doi:10.1037/0033-2909.125.6.677

Elliot, A., Faler, J., McGregor, H., Campbell, W., Sedekides, C., & Harackiewicz, J. (2000). Competence valuation as a strategic intrinsic motivation process. *Personality and Social Psychology Bulletin, 26,* 780–794. doi:10.1177/0146167200269004

Elliot, A. J., & Moller, A. C. (2003). Performance-approach goals: Good or bad forms of regulation? *International Journal of Educational Research, 39,* 339–356. doi:10.1016/j.ijer.2004.06.003

Ericsson, K. A., Krampe, R. Th., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100,* 363–406. doi:10.1037/0033-295X.100.3.363

Grolnick, W., & Ryan, R. M. (1987). Autonomy in children's learning: An experimental and individual difference investigation. *Journal of Personality and Social Psychology, 52,* 890–898. doi:10.1037/0022-3514.52.5.890

Harackiewicz, J. M. (1979). The effects of reward contingency and performance feedback on intrinsic motivation. *Journal of Personality and Social Psychology, 37,* 1352–1363. doi:10.1037/0022-3514.37.8.1352

Harackiewicz, J. M., Abrahams, S., & Wageman, R. (1987). Performance evaluation and intrinsic motivation: The effects of evaluative focus, rewards, and achievement orientation. *Journal of Personality and Social Psychology, 53,* 1015–1023. doi:10.1037/0022-3514.53.6.1015

Harackiewicz, J. M., Manderlink, G., & Sansone, C. (1984). Rewarding pinball wizardry: Effects of evaluation and cue value on intrinsic interest. *Journal of Personality and Social Psychology, 47,* 287–300. doi:10.1037/0022-3514.47.2.287

Harter, S. (1978). Pleasure derived from challenge and the effects of receiving grades on children's difficulty level choices. *Child Development, 49,* 788–799. doi:10.2307/1128249

Hayes, A. F. (2011). *Indirect.* Retrieved from http://www.afhayes.com/public/indirect.pdf

Hayes, A. F. (2012). *Process.* Retrieved from http://www.afhayes.com/public/process.pdf

Hayes, A. F., & Preacher, K. J. (2011). *Indirect and direct effects of a multicategorical causal agent in statistical mediation analysis.* Manuscript submitted for publication.

Henin, J., Accorsi, E., Cho, P., & Tabor, W. (2009). Extraordinary natural ability: Anagram solution as an extension of normal reading ability. In the *Proceedings of the 31st Annual Meeting of the Cognitive Science Society.* Mahwah, NJ: Lawrence Erlbaum Associates.

Houlfort, N., Koestner, R., Joussemet, M., Nantel-Vivier, A., & Lekes, N. (2002). The impact of performance-contingent rewards on task autonomy and competence. *Motivation and Emotion, 26,* 279–295. doi:10.1023/A:1022819120237

Jarque, C. M., & Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistics Review, 55,* 163–172. doi:10.2307/1403192

Kenny, D. A. (2011). *Path analysis.* Retrieved from http://davidakenny.net/cm/pathanal.htm

Lent, R. W., Brown, S. D., & Larkin, K. C. (1987). Comparison of three theoretically derived variables in predicting career and academic behavior: Self-efficacy, interest congruence, and consequence thinking. *Journal of Counseling Psychology, 34,* 293–298. doi:10.1037/0022-0167.34.3.293

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,* 129–137. doi:10.1037/h0035519

Liu, W. C., Wang, C. K. J., & Parkins, E. J. (2005). A longitudinal study of students' academic self-concept in a streamed setting: The Singapore context. *British Journal of Educational Psychology, 75,* 567–586.

Lust, J. (2004). Judy Cameron and W. David Pierce: Rewards and intrinsic motivation: Resolving the controversy (Book review). *Personnel Psychology, 57,* 259–261.

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis.* Mahwah, NJ: Erlbaum.

Maehr, M. L., & Stallings, W. M. (1972). Freedom from external evaluation. *Child Development, 43,* 177–185. doi:10.2307/1127881

Mardia, K. V. (1985). Mardia's test of multinormality. In S. Kotz & N. L. Johnson (Eds.), *Encyclopedia of statistical sciences* (Vol. 5, pp. 217–221). New York, NY: Wiley.

Midgley, C., Kaplan, A., & Middleton, M. (2001). Performance-approach goals: Good for what, for whom, under what circumstances, and at what cost? *Journal of Educational Psychology, 93,* 77–86. doi:10.1037/0022-0663.93.1.77

Novick, L. R., & Sherman, S. J. (2003). On the nature of insight solutions: Evidence from skill differences in anagram solution. *The Quarterly Journal of Experimental Psychology, 56*A, 351–382.

Nygreen, G. T. (1971). Interactive path analysis. *The American Sociologist, 6,* 37–43.

Oakes, J. 1985. *Keeping track: How schools structure inequality.* New Haven, CT. Yale.

Osborne, J. (2002). Notes on the use of data transformations. *Practical Assessment, Research & Evaluation, 8*(6). Retrieved from http://PAREonline.net/getvn.asp?v=8&n=6

Patrick, B. C., Skinner, E. A., & Connell, J. P. (1993). What motivates children's behavior and emotion? Joint effects of perceived control and autonomy in the academic domain. *Journal of Personality and Social Psychology, 65,* 781–791. doi:10.1037/0022-3514.65.4.781

Pierce, T., & Vaca, R. (2007). Distracted: Academic performance differences between teen users of MySpace and other communication technologies. *Proceedings EISTA.* Retrieved from http://www.iiisci.org/journal/CV$/sci/pdfs/E214BL.pdf

Pierce, W. D., Cameron, J., Banko, K. M., & So, S. (2003). Positive effects

of rewards and performance standards on intrinsic motivation. *The Psychological Record, 53,* 561–579.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40,* 879–891. doi:10.3758/BRM.40.3.879

Pritchard, R. D., Campbell, K. M., & Campbell, D. J. (1977). Effects of extrinsic financial rewards on intrinsic motivation. *Journal of Applied Psychology, 62,* 9–15. doi:10.1037/0021-9010.62.1.9

Pulfrey, C., Buchs, C., & Butera, F. (2011). Why grades engender performance-avoidance goals: The mediating role of autonomous motivation. *Journal of Educational Psychology, 103,* 683–700. doi:10.1037/a0023911

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology, 43,* 450–461. doi:10.1037/0022-3514.43.3.450

Ryan, R. M., & Connell, J. P. (1989). Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology, 57,* 749–761. doi:10.1037/0022-3514.57.5.749

Ryan, R. M., & Deci, E. L. (1989). Bridging the traditions of task/ego involvement and intrinsic/extrinsic motivation: A comment on Butler (1987). *Journal of Educational Psychology, 81,* 265–268. doi:10.1037/0022-0663.81.2.265

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25,* 54–67. doi:10.1006/ceps.1999.1020

Ryan, R. M., & Deci, E. L. (2006). Self-regulation and the problem of human autonomy: Does psychology need choice, self-determination, and will? *Journal of Personality, 74,* 1557–1586. doi:10.1111/j.1467-6494.2006.00420.x

Ryan, R. M., Koestner, R., & Deci, E. L. (1991). Varied forms of persistence: When free-choice behavior is not intrinsically motivated. *Motivation and Emotion, 15,* 185–205. doi:10.1007/BF00995170

Ryan, R. M., & Weinstein, R. M. (2009). Undermining quality teaching and learning: A self-determination theory perspective on high-stakes testing. *Theory and Research in Education, 7,* 224–233. doi:10.1177/1477878509104327

Schunk, D. (1981). Modeling and attributional feedback effects on children's achievement: A self-efficacy analysis. *Journal of Educational Psychology, 73,* 93–105. doi:10.1037/0022-0663.73.1.93

Schunk, D., & Hanson, A. R. (1985). Peer models: Influence on children's self-efficacy and achievement behaviors. *Journal of Educational Psychology, 77,* 313–322. doi:10.1037/0022-0663.77.3.313

Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis.* New York, NY: Appleton-Century-Crofts.

Slavin, R. E. (1988). Synthesis of research on grouping in elementary and secondary schools. *Educational Leadership, 46,* 67–77.

Srivastava, M. S. (1984). A measure of skewness and kurtosis and a graphical method for assessing multivariate normality. *Statistics & Probability Letters, 2,* 263–267. doi:10.1016/0167-7152(84)90062-2

Stapel, D. A., & Koomen, W. (2005). Competition, cooperation, and the effects of others on me. *Journal of Personality and Social Psychology, 88,* 1029–1038. doi:10.1037/0022-3514.88.6.1029

Toma, C., & Butera, F. (2009). Hidden profiles and concealed information: Strategic information sharing and use in group decision making. *Personality and Social Psychology Bulletin, 35,* 793–806. doi:10.1177/0146167209333176

Trautwein, U., Ludtke, O., Marsh, H. W., Köller, O., & Baumert, J. (2006). Tracking, grading, and student motivation: Using group composition and status to predict self-concept and interest in ninth-grade mathematics. *Journal of Educational Psychology, 98,* 788–806. doi:10.1037/0022-0663.98.4.788

20 Minutes Online. (2010, November 24). *La filiére VSO n'est pas discriminatoire* [The VSO course of study is not discriminating]. Retrieved February 10, 2012, from http://www.20min.ch/ro/news/vaud/story/20175962/

Vanlede, M., Philippot, P., & Galand, B. (2006). Croire en soi: Le rôle de la mémoire autobiographique dans la construction du sentiment d'efficacité [Believing in oneself: The role of autobiographical memory in self-efficacy construction]. In B. Galand & E. Bourgeois (Eds.), *(Se) Motiver à Apprendre* (pp. 51–54). Paris: Presses Universitaires de France.

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review, 66,* 297–333.

Wicker, F. W., Brown, G., & Paredes, V. (1990). Competing activities and measures of intrinsic motivation. *Journal of Social Psychology, 130,* 813–819. doi:10.1080/00224545.1990.9924633

Wicker, F. W., Brown, G., Wiehe, J. A., & Shim, W. Y. (1990). Moods, goals and measures of intrinsic motivation. *The Journal of Psychology, 124,* 75–86.

Wolf, P. J. (2007). Academic improvement through regular assessment. *Peabody Journal of Education, 82,* 690–702. doi:10.1080/01619560701603114

Zhang, W., Liu, M., Zhao, S., & Xie, Q. (2011). English test-taking strategy use and students' test performance. *Asian EFL Journal, 13,* 133–168.

Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal, 29,* 663–676.