UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Promoting performance and motivation through a combination of intrinsic motivation stimulation and an extrinsic incentive

Wijsman, L.A.; Saab, N.; Schuitema, J.; van Driel, J.H.; Westenberg, P.M.

DOI
10.1007/s10984-018-9267-z

Publication date
2019

Document Version
Final published version

Published in
Learning Environments Research

License
CC BY

Link to publication

Citation for published version (APA):
Wijsman, L. A., Saab, N., Schuitema, J., van Driel, J. H., & Westenberg, P. M. (2019). Promoting performance and motivation through a combination of intrinsic motivation stimulation and an extrinsic incentive. *Learning Environments Research*, *22*(1), 65-81. https://doi.org/10.1007/s10984-018-9267-z

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:05 Sep 2026

Learning Environ Res (2019) 22:65–81
https://doi.org/10.1007/s10984-018-9267-z

ORIGINAL PAPER

Promoting performance and motivation through a combination of intrinsic motivation stimulation and an extrinsic incentive

Lindy A. Wijsman[1] · Nadira Saab[2] · Jaap Schuitema[3] · Jan H. van Driel[4] · P. Michiel Westenberg[5]

Received: 11 July 2017 / Accepted: 18 April 2018 / Published online: 30 April 2018
© The Author(s) 2018

Abstract During the early stages of secondary education students' motivation and performance levels decline. This study employed a case study approach to evaluate a learning environment called GUTS; Differentiated Challenging of Talent in School. GUTS was specifically designed to raise performance and motivation through a combination of (a) talent lessons as an intrinsic motivator and (b) a higher promotion standard as an extrinsic incentive. Participants were 156 students who started secondary education in grade 7 in school year 2013/2014, and participated in GUTS for three successive school years. Performance and motivation measures were longitudinally collected between grades 7 and 9 and were analysed in comparison to previous cohorts of students at the same school and students at other schools. Additionally, measures of well-being and self-esteem were included to explore possible collateral effects of the intervention. The GUTS cohort displayed generally higher levels of performance and motivation than the comparison groups. However, the GUTS cohort still showed the decline of performance and motivation between grades 7 and 9 that was also observed in the comparison groups. This study showed that performance and motivation levels were increased without collateral damage to the students' overall sense of well-being and self-esteem, while GUTS was not a strong enough intervention to counter the motivation and performance decline over time. Possible reasons for the persistent decline of student performance and motivation are discussed in terms of

✉ Lindy A. Wijsman
l.a.wijsman@fsw.leidenuniv.nl

[1] Unit Developmental and Educational Psychology, Institute of Psychology, Leiden University, Wassenaarseweg 52, 2333 AK Leiden, The Netherlands

[2] ICLON, Leiden University, Kolffpad 1, 2333 BN Leiden, The Netherlands

[3] Educational Sciences, University of Amsterdam, Nieuwe Achtergracht 127, 1001 NG Amsterdam, The Netherlands

[4] Melbourne Graduate School of Education, The University of Melbourne, 234 Queensberry Street, Melbourne, VIC 3010, Australia

[5] Unit Developmental and Educational Psychology, Institute of Psychology, Leiden University, Wassenaarseweg 52, 2333 AK Leiden, The Netherlands

various factors at the level of GUTS, the educational context and the needs of the developing adolescent.

Keywords Extrinsic incentive · GUTS · Intrinsic motivation stimulation · Motivation · Performance

Introduction

During the early stages of secondary education, student motivation and performance levels decline (Eccles et al. 1991; Peetsma et al. 2005; Wijsman et al. 2016). As this appears to occur in several countries, it is not restricted to a particular education system. Declining motivation and performance are problematic as both factors are important determinants of students' school success. We used a case-study approach to evaluate the effect of a learning environment specifically designed to raise motivation *and* performance.

Decreasing motivation and performance in lower secondary education

The literature mentions three factors that contribute to decreasing motivation and performance in lower secondary education. Firstly, a student factor that explains low motivation in secondary school is that it is not cool to show to peers that you are putting effort into your schoolwork (Mijs and Paulle 2016; Warrington et al. 2000). The willingness of students to portray themselves as diligent, especially towards peers, decreases as students come of age (Juvonen and Murdock 1995).

Secondly, the lack of motivation in school has been assigned to a mismatch between adolescents' needs and what schools offer (Eccles et al. 1993) and to shifting relevance of life domains for adolescents (Peetsma 1997; as cited in Peetsma and van der Veen 2011). The mismatch between what students expect from school and what they experience appears to be connected to their affective state. Könings et al. (2011) showed that motivational and concentration problems are related to a mismatch between students' preferences and perceptions of classroom characteristics. In these explanations, the underlying mechanism is the failure to meet students' needs.

Thirdly, education professionals in several countries have recognised that the education system does not encourage students to reach their academic potential. Students in this context do what is required instead of meeting their own full potential. This tendency has been referred to as a culture of C's. In several countries, the education system and school policy have attempted innovations to tackle this culture of C's and stimulate a more ambitious learning culture (Danish Ministry of Education 2014; De Boer et al. 2013; Ministry of Education and Culture of Finland 2014; Dutch Ministry of Education Culture and Science 2011).

Dutch context

In the Netherlands, the culture of C's might be even more prominent than in other countries. While Dutch students generally perform well in international comparisons, the proportion of top performers is not as high as in Asian countries (OECD 2016). The OECD (2016, p. 79) wrote that "some of the most promising students in the Netherlands are not

reaching their full potential". In addition, student motivation is generally low in comparison with other countries (OECD 2016).

A specific reason for this culture of C's in The Netherlands is the way in which students progress through secondary education. To progress from one grade to the next, students need to meet a set of specific criteria. Schools have some freedom to determine the specific criteria but, in each school, the bottom line is defined by the lowest allowable grade, and that is 6 on a scale of 1 (extremely poor) to 10 (perfect). Overall, students need to obtain at least a 6 for their subjects to progress to the next grade. At the same time, performances higher than this minimum standard do not have consequences for students' progress in school. This might create a focus on the lowest possible grade rather than aiming for higher grades, thus allowing students to lower their efforts and create more free time for their hobbies and friends. Indeed, a recent study in the Dutch context shows that average report card grades (GPA) gradually decrease during lower secondary education (Wijsman et al. 2016). Our research was the first longitudinal study to investigate GPA of a large group of students between grades 7 and 9, and it showed that students in seventh grade achieve a GPA above 7 and that, in the following two grades, these GPA levels linearly decline in the direction of an average 6.

Combining intrinsic and extrinsic motivators to stimulate performance and motivation

To improve student performance and motivation in the Dutch context, we designed an innovation based on a combination of extrinsic and intrinsic motivators. A meta-analysis of Cerasoli et al. (2014) showed that a combination of intrinsic motivation and extrinsic incentives is important for performance. Intrinsic motivation predicts quality of performance best, whereas extrinsic incentives are the best predictors of quantity of performance. It was also found that the presence of extrinsic incentives boosted the link between intrinsic motivation and performance (Cerasoli et al. 2014).

At the same time, extrinsic incentives closely related to performance could fuel a feeling of being controlled and pressured, which might cause collateral damage effects to one's individual sense of well-being (Cerasoli et al. 2014). Cognitive Evaluation Theory (CET), as part of self-determination theory (Ryan and Deci 2002), postulates that incentives negatively influence self-determination through pressure from outside the individual to perform an activity if they diminish one's basic psychological need for autonomy. At the same time, incentives are theorised to provide information about one's competence, and this too is one of the basic psychological needs to be fulfilled to stimulate intrinsic motivation. Whether incentives then negatively or positively relate to intrinsic motivation depends on the strength of the detrimental effect on autonomy compared with the advantageous effect on competence.

For the present study, we made an attempt to capitalise on the combined effects of intrinsic and extrinsic factors, while avoiding potentially aversive effects of an extrinsic incentive. We created a learning environment in which positive expectations for the students' overall performance level were coupled with a specific focus on the students' strengths and personal interests.

Learning environment

The learning environment evaluated in the present study is called GUTS (in Dutch: Gedifferentieerd Uitdagen van Talent op School; *Differentiated Challenging of Talent at School*). The overall goal is to increase performance level *and* motivation for school. GUTS combines two elements to stimulate performance and motivation: (1) talent lessons as an intrinsic motivator and (2) higher standards for promotion as an extrinsic incentive.

Talent lessons were provided by means of extra lessons in specific subjects. Students chose subjects they liked and were good at. The goal of the talent lessons was to stimulate and continue students' intrinsic motivation for that subject, and performance and motivation for school in general, through positively approaching students in a subject that they liked. The talent lessons in the present study focused on students' strengths and aimed to fulfil students' need for autonomy. When feeling autonomous, one's behaviour is perceived in accordance with one's interests and values, and activities are experienced as volitional (Ryan and Deci 2002). In a school setting, the need for autonomy can be fulfilled through autonomy-supportive teaching, which means that teachers incorporate students' interests and values through offering choice, fostering relevance and showing respect (Stroet et al. 2013; Vansteenkiste et al. 2012). If students perceive autonomy-support, they experience more motivation and engagement in school (Reeve 2006; Stroet et al. 2013).

A higher promotion standard was implemented as an extrinsic incentive. In GUTS, a report card grade of 6 for a single subject was still sufficient, but students needed to achieve an average of 7 at the end of the school year. Report card grades are measures of performance at various tests during a school year. The higher promotion standard was thought to change the behaviour of students is different ways. Firstly, by asking more from students, it becomes worthwhile to perform at a higher level. Most students only put effort into an activity when they think that this effort is valuable (Wigfield and Eccles 2000). As Warrington et al. (2000) illustrated, students deliberately choose the subjects at which they want to excel and for which a pass is enough. Secondly, students respond differently to low and high grades. Whereas low grades strengthen the decline of involvement with school, high grades reduce this decline in involvement (Poorthuis 2012). Thirdly, the promotion standard can affect the culture of C's among peers in schools. If everyone excels in some subjects and not in others, then it becomes normal for everyone to excel on occasion and this can lessen the 'uncool' stigma.

Research questions

To examine student motivation and performance in the context of GUTS, this study employed a case study approach. A case study is suitable when investigating a phenomenon that cannot be separated from its context (Yin 2011). GUTS was implemented at a Dutch secondary school and was piloted for three subsequent years. Students were followed from seventh to ninth grade. To take into account the possibility of collateral damage triggered by GUTS, we also assessed the students' sense of well-being and self-esteem. Two research questions were addressed.

Question 1 What is the effect of GUTS on student performance? This was evaluated by comparing the GUTS group with: (a) a cohort of other schools and (b) previous cohorts at the same school. We expected that a combination of higher performance expectations (extrinsic incentive) with talent lessons (intrinsic motivation stimulation) would lead to better performance in the GUTS cohort

Question 2 What is the effect of GUTS on student motivation, well-being and self-esteem? This was evaluated by comparing the GUTS cohort with national norm data. We expected that a combination of higher performance expectations with explicit attention to students' basic psychological need-fulfilment in the talent lessons would improve motivation and not undermine the students' sense of well-being and self-esteem.

Method

Participants

GUTS was implemented at a bilingual secondary school in the western part of the Netherlands. This school offered higher general secondary education and pre-university tracks. A tracking system is used in Dutch secondary schools, placing students in a track before entering grade 7. The two highest tracks were included in the present study. One subgroup of the participants was in the higher general secondary track. This school type prepares students for higher professional education. The majority of students was in the pre-university track. This school type prepares students for university. In seventh grade, no separate general secondary track exists in the participating school. Instead, a combined higher general secondary/pre-university track is offered.

Participants consisted of 184 students who were divided into eight grade 7 classes and who started their secondary education in school year 2013/2014. Twenty-eight students could not be included in the data analyses for the present study because seven students had left the school between grade 7 and 9 for personal reasons (e.g. family moved away), 13 students left the school because of insufficient grades, and eight students had to repeat a grade. A total of 21 out of 184 students (11.4%) were excluded from data analyses because of poor grades. This percentage does not differ significantly from grade-related dropout in previous cohorts at the same school before the implementation of GUTS (cohort − 1: 7.4%; $\chi^2 = 1.60$, $df = 1$, $p = 0.206$, and cohort 0: 9.7%;

Table 1 GUTS cohort: number of students in both school types and all school years

School type	Sample size		
	2013/2014 Grade 7	2014/2015 Grade 8	2015/2016 Grade 9
General secondary (/ pre-university)	70	44	34
Pre-university	114	126	122
Total	184	170	156

$\chi^2 = 0.27$, $df = 1$, $p = 0.601$). The final sample consisted of 156 participants who were in school from seventh to ninth grade (see Table 1).

Informed consent was provided by the parents or caretakers of the students involved. In total, for 18 out of 202 students, no consent was provided for participation in the studies. This procedure was validated by the first author's institution's Research Ethics Committee.

Materials

Learning environment

GUTS combined talent lessons and a higher standard for promotion. *Talent lessons* were designed to fulfil the need for autonomy in multiple ways. First, choosing a subject for the talent lessons offered some choice to the students to spend extra time on one of their interests. For example, if physics was a student's preferred subject, he or she was allowed to follow talent lessons on this subject. Second, during the talent lessons, as much freedom as possible was offered to choose and structure one's own activities. Moreover, the teacher differentiated between the students, by taking differences between students into account (Tomlinson et al. 2003). This provides students with positive affect and motivation towards learning (Tomlinson et al. 2003). Finally, higher-order thinking tasks and enrichment were also included as principles in the talent lessons as aspects that challenge and interest students and make them acquainted with topics that are not included in the standard curriculum.

In grade 7 and 8, two cycles of eight two-hour lessons were provided. Each school year, students could choose two different subjects for the talent lessons. Additionally, in eighth grade, presentations for parents at the end of a talent lessons cycle were initiated. In ninth grade, students completed a personal project whose goal was to stimulate the student's intrinsic motivation for the subject. Students individually worked on a topic that they deemed interesting and they had to deliver a tangible product. The personal project was autonomy-supportive based on the same principles as the talent lessons. Students could choose their own topic for the project and, during the process, they had much freedom within the boundaries of the project. Students worked together in groups of 'critical friends' to make a plan for the project, which had to be approved by the supervising teacher.

The higher *performance standard* included that students needed to achieve an average report card grade of at least 7 at the end of every school year. This means that if they achieved lower-than-7 report card grades in some subjects, this had to be compensated by obtaining higher-than-7 report card grades for other subjects. In this way, students could differentiate their achievement between subjects, giving everyone the opportunity to excel in some subjects and compensate for other subjects that were of less interest to them.

Instruments

Student performance was conceptualised as report card grades for all subjects in seventh, eighth and ninth grade. Grades theoretically range between 1 (extremely poor) and 10 (perfect). Students receive a grade report at the end of every school year, which contains the report card grades for all subjects that the student took in that particular year. For this

study, an average of all subjects was calculated for every student. We have analysed average report card grades in previous research on performance patterns because of the importance of grades for student school careers (Wijsman et al. 2016).

Student motivation for school, sense of well-being and self-esteem were assessed with the School Attitude Questionnaire (SAQ; Vorst et al. 2010). Reliability of the three scales in previous research ranged between $\alpha = 0.84$ and $\alpha = 0.90$ (Vorst et al. 2010). Each scale consists of three subscales with eight items. *Motivation for school* consisted of the subscales of learning task orientation, concentration in the class, and homework attitude. *Well-being at school* consists of fun at school, feeling socially accepted and relating to teachers. Finally, *self-esteem* consists of expression skills, confidence at tests, and social skills. The social desirability scale was not included in the present study. Students answered the items on a three-point scale (agree, don't know, don't agree). Potentially, scores on the scales range between a minimum of 24 and a maximum of 72. The SAQ questionnaire is a broadly accepted diagnostic tool in the Netherlands and has been used in research (Könings et al. 2005; Lee et al. 2012). The scoring manual includes extensive norm data, which allows for comparison with the data from our GUTS cohort.

Procedure

To design GUTS in accordance with school policy and procedures, researchers from Leiden University regularly met with the school management team to discuss the development and implementation of GUTS. In addition, the researchers provided several presentations for teachers and parents of students to increase their understanding and appreciation of this innovation at the school. Feedback from teachers and students in the initial phase of the project was also used to fine-tune GUTS.

The higher performance norm was initiated from the start of the project in grade 7. Discussed explicitly with the school was that grading of examinations and other assignments would be provided based on the same standards as in previous years. The specific design of the talent lessons developed as the project proceeded, although the idea of positive attention for a subject one liked and was good at was embedded from the beginning.

Data collection was spread over three years and occurred every year in October and June. The SAQ was administered in October of each year (grade 7, 8 and 9) by the first author in a regular classroom setting. The teachers were present in the classroom and completion of the questionnaire took approximately 30 min. In June of every school year from grade 7 to grade 9, students' report card grades were collected from the school's student tracking system.

Analyses

The effect of GUTS on student performance, motivation, well-being and self-esteem was studied by comparing the GUTS cohort with non-GUTS cohorts. For each measure, the most appropriate available comparison group was chosen.

For Research question 1, the effect of GUTS on student performance was studied by comparing the GUTS cohort with two non-GUTS comparison groups. The first comparison group consisted of a large sample of students ($n = 667$) from five high schools including two school types – higher general secondary and pre-university education – that participated in a previous study (Wijsman et al. 2016). The second comparison group consisted

of students in the same school in two preceding cohorts: school year 2011/2012 (cohort −1; $n=132$) and school year 2012/2013 (cohort 0; $n=153$).

The effect of GUTS on performance was studied using hierarchical regression analyses. Two levels were distinguished in all analyses, time point within student (level 1) and student (level 2). Time point within student refers to the multiple measurements that were included for each student. Students who are in the same class are taught by the same teacher, which can make them more alike then students across classes. To consider this nesting of students in classes, a third level was added in the models comparing the GUTS cohort to previous cohorts. Because some classrooms were rearranged during the three years, we formed clusters of students who were in the same class between grades 7 and 9. In the comparison with other schools, no clusters were formed because classes were not known for all students. Several piecewise linear models were fitted. Model 1 in both comparisons served as the variance component model, and included two time predictors: the difference between grade 7 and 8 (t2) and the difference between grade 8 and 9 (t3). The factor time was included in the first model to prevent an overestimation of the variance at the student level (Hox 2010). Model 2 either included as a predictor the group of comparison schools (No GUTS) or the previous cohorts of the case school (cohorts −1 and 0), and the interactions with time. In the following model, various control variables were included, depending on the available measures. These were school type (higher general secondary or pre-university), sex and the so-called Cito-score. The latter is the score on a standardised national test measuring academic aptitude, which is administered at the end of grade 6 (primary school), to help determine the best fitting school type for students. We also controlled for students who changed school types between grades 7 and 9. Changes in school type affect report card grades, because the level of difficulty differs between school types. We therefore included two dummy variables for students who moved up (from higher general secondary to pre-university) and for students who moved down, respectively. Additionally, separate analyses were performed for the difference in performance level between each comparison group and GUTS students in grades 7 to 9 separately. In the comparison with other schools, ANOVAs were used and, in the comparison with previous cohorts at the same school, multilevel regression analyses were used because of the variance located at the cluster level.

For Research question 2 involving the effect of GUTS on motivation, self-esteem and well-being, students at the GUTS school were compared to a national norm group included in the SAQ manual (Vorst et al. 2010). For this cross-sectional norm group, mean scores, standard deviations and sample sizes were available. For each grade, independent *t* tests were used to compare these cross-sectional data in the SAQ manual with the longitudinal data from our GUTS cohort. Repeated-measures analyses were not possible because we did not have access to scale or item ratings for each student in the SAQ norm group. At first, sex was included in the analyses but, because no interaction effects were observed, sex was excluded in the final analyses presented below. For students in the GUTS school, some missing data were present. The three SAQ-scales were constructed by taking the total sum scores of the items and, consequently, a missing score on only one scale item would result in a missing scale score. Therefore, missing values on the items of the three scales were imputed using Expectation Maximization (Tabachnick and Fidell 2007). In grade 7, maximally 1.1% of the data were missed per item, in grade 8, maximally 0.6% of the data missed per item and, in grade 9, maximally 5% of the data were missed per item.

Results

Research question 1: the effect of GUTS on performance

Mean report card grades for the GUTS and non-GUTS comparison cohorts are presented in Table 2 and are plotted in Figs. 1 and 2. The data show a consistent pattern across both comparisons in that (a) the GUTS cohort performed better than all comparison cohorts and in all three grades and (b) the level of performance decreased between grade 7 and 9 for all student cohorts, including the GUTS cohort. In other words, GUTS appeared to have had a main effect on performance level, but GUTS did not change the degree of the performance decline in lower secondary education.

Hierarchical linear regression analyses were performed to test the comparison between the GUTS and the non-GUTS schools between grade 7 and 9 (see Table 3). The final model (Model 3 in Table 3) showed that overall performance levels for the groups together significantly decreased in time ($b=-0.28$, $p<0.001$ between grade 7 and grade 8, and $b=-0.36$, $p<0.001$ between grade 8 and grade 9). Also, the performance of students in the comparison group was significantly lower ($b=-0.15$, $p<0.001$) than that of students in the GUTS group in grade 7. Between grade 7 and 8, the non-GUTS students' performance decreased significantly more than that of students in the GUTS group ($b=-0.09$, $p=0.004$), and between grade 8 and 9, unexpectedly, the performance of students in the GUTS group decreased significantly more ($b=0.17$, $p<0.001$)). Figure 1 presents performance means of the two groups in grade 7, 8 and 9. Even with this decline in performance of the GUTS-students, the difference in report card grades between the group of students without GUTS and with GUTS was significant in all three years (grade 7: $F(1,822)=26.04$, $p<0.001$; grade 8: $F(1,822)=46.35$, $p<0.001$; grade 9: $F(1,822)=11.99$, $p=0.001$). To control for the possibility that this GUTS versus Non-GUTS difference was due in part to students who switched between the school types, we studied potential interaction effects for students who 'moved up' from secondary to pre-university and for student who 'moved down' from pre-university to general. A significant interaction was found between time and 'moving up' (grade 8: $b=0.10$, $p=0.002$; grade 9: $b=-0.27$, $p<0.001$). This means that the performance of students who changed to a higher school type decreased less between grade 7 and 8 and more between grade 8 and 9 in comparison to students who stayed in the same school type.

Table 2 Average report card grades of the various groups in grades 7 to 9 in comparison and GUTS groups

Group	*N*	Grade 7		Grade 8		Grade 9	
		M	*SD*	*M*	*SD*	*M*	*SD*
Comparison group (no GUTS)	667	7.24	0.52	6.92	0.53	6.64	0.59
Case school							
Cohort − 1 (no GUTS)	132	7.22	0.56	7.02	0.66	6.74	0.74
Cohort 0 (no GUTS)	153	7.10	0.55	6.94	0.56	6.61	0.69
Cohort 1 (GUTS)	156	7.48	0.51	7.23	0.47	6.82	0.66

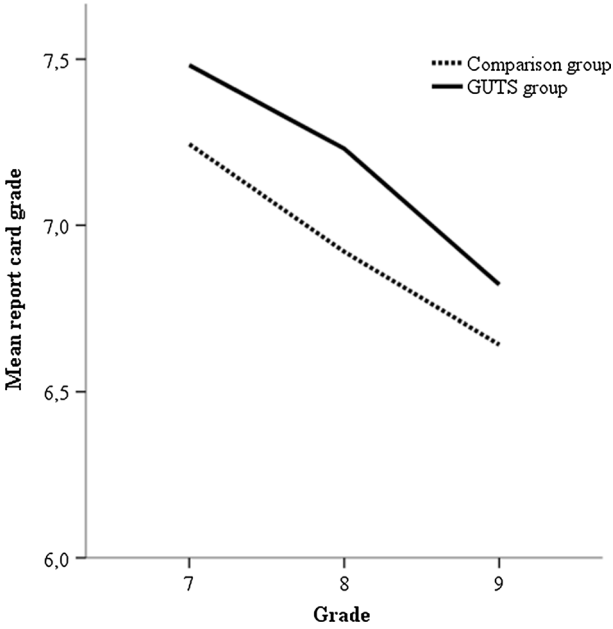

Fig. 1 Performance means of the GUTS group and the comparison group

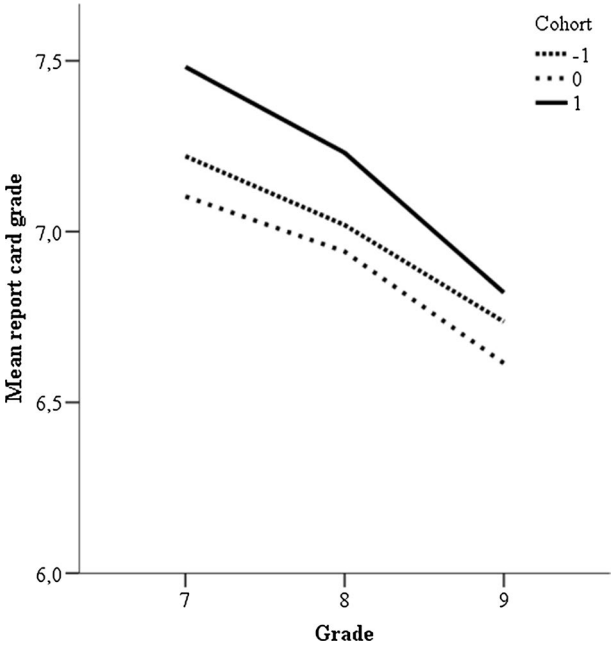

Fig. 2 Performance means of the GUTS cohort (1) and two previous cohorts without GUTS (cohort − 1, and 0)

Table 3 Performance growth model of performance for GUTS students and comparison group at other schools

Effects	Model 1		Model 2		Model 3	
	b	*SE*	*b*	*SE*	*b*	*SE*
Fixed effects						
Intercept	7.29***	0.02	7.48***	0.04	7.50***	0.04
Level 1 (*time*)						
Time2	−0.31***	0.01	−0.25***	0.03	−0.28***	0.03
Time3	−0.30***	0.01	−0.41***	0.03	−0.36***	0.04
Level 2 (*participant*)						
No GUTS			−0.24***	0.05	−0.15***	0.04
T2 × No GUTS			−0.07*	0.03	−0.09**	0.03
T3 × No GUTS			0.13***	0.04	0.17***	0.04
Gen.secondary					−0.78***	0.14
T2 × Gen.secondary					0.17	0.11
T3 × Gen.secondary					0.05	0.13
Girls					0.21***	0.03
T2 × Girls					0.03	0.02
T3 × Girls					−0.00	0.03
Up					−0.03	0.04
Down					0.25	0.14
T2 × Up					0.10**	0.03
T3 × Up					−0.27***	0.04
T2 × Down					−0.15	0.11
T3 × Down					−0.08	0.13
Random effects						
Variance time 1	0.28	0.01	0.27	0.01	0.19	0.01
Variance time 2	0.29	0.01	0.27	0.01	0.19	0.01
Variance time 3	0.37	0.02	0.37	0.02	0.30	0.02
Covariance 2,1	0.22	0.01	0.21	0.01	0.13	0.01
Covariance 3,1	0.20	0.01	0.20	0.01	0.13	0.01
Covariance 3,2	0.24	0.01	0.24	0.01	0.17	0.01
−2LogLikelihood	2625.08		2575.63		2193.19	

*$p < 0.05$, two-tailed; **$p < 0.01$, two-tailed; ***$p < 0.001$, two-tailed

Hierarchical linear regression analyses were performed to test the comparison between the GUTS cohort and the two pre-GUTS cohorts at the same school (see Table 4). In total, 69 clusters of students who were in the same class in all three grades were distinguished and included in the analyses. The final model shows that, overall, performance levels decreased for all students. Model 3 in Table 4 shows that, in grade 7, the differences between GUTS and the two no GUTS groups were significant (cohort −1: $b = -0.26$, $p < 0.001$; cohort 0: $b = -0.31$, $p < 0.001$). Furthermore, the development of performance between grade 7 and 9 of students in cohort −1 and cohort 0 was not significantly different from the development of the performance from the GUTS cohort.

Table 4 Performance growth model of performance for GUTS students and students in two previous cohorts at the same school

Effects	Model 1 $n=441$		Model 2 $n=441$		Model 3 $n=372$	
	b	*SE*	*b*	*SE*	*b*	*SE*
Fixed effects						
Intercept	7.22***	0.05	7.45***	0.07	6.95***	0.07
Level 1 (*time*)						
Time2	−0.21***	0.02	−0.25***	0.03	−0.17**	0.05
Time3	−0.34***	0.02	−0.41***	0.04	−0.24***	0.06
Level 2 (*participant*)						
Cohort − 1			−0.39**	0.12	−0.26***	0.07
Cohort 0			−0.40**	0.11	−0.31***	0.06
T2 × cohort − 1			0.05	0.04	0.05	0.05
T3 × cohort − 1			0.13**	0.05	0.11	0.06
T2 × cohort 0			0.09*	0.04	0.09	0.05
T3 × cohort 0			0.08	0.05	0.06	0.06
Citoscore					0.05	0.03
T2 × citoscore					0.04	0.02
T3 × citoscore					0.06	0.03
Pre-univ.					0.49***	0.07
T2* Pre-univ.					−0.12*	0.05
T3* Pre-univ.					−0.15*	0.06
Girl					0.29***	0.05
T2 × Girl					−0.00	0.04
T3 × Girl					−0.11*	0.05
Random effects						
Variance time 1	0.22	0.02	0.22	0.02	0.19	0.01
Variance time 2	0.26	0.02	0.25	0.02	0.23	0.02
Variance time 3	0.44	0.03	0.43	0.03	0.40	0.03
Covariance 2,1	0.17	0.02	0.17	0.02	0.14	0.01
Covariance 3,1	0.17	0.02	0.18	0.02	0.15	0.02
Covariance 3,2	0.24	0.02	0.24	0.02	0.22	0.02
Variance cluster	0.11	0.03	0.08	0.03	0.01	0.01
− 2LogLikelihood	1548.50		1522.96		1170.45	

*$p<0.05$, two-tailed; **$p<0.01$, two-tailed; ***$p<0.001$, two-tailed

For the differences between GUTS and no GUTS in grades 8 and 9, two separate multilevel linear regression analyses were performed at these two time points, including all predictors. Significant differences were also found in grade 8 (cohort − 1: $b=-0.24$, $p=0.004$; cohort 0: $b=-0.26$, $p=0.001$). In grade 9, no variance was located at the cluster level after adding all predictors to this model. Therefore, the cluster level was removed from the analysis. No significant differences between GUTS and cohort − 1 was found ($b=-0.09$, $p=0.30$) and a slight difference with cohort 0 was found ($b=-0.16$, $p=0.047$).

Research question 2: level and development of students' motivation for school, well-being and self-esteem

To investigate students' motivation for school, well-being and self-esteem during grades 7 to 9, a comparison was made with a national norm group as described in Vorst et al. (2010). Table 5 provides results for the differences between GUTS and the norm group. A consistent pattern was observed for motivation: students at the GUTS school reported significantly higher motivation than the students in the norm group and this occurred across all three grades. In addition, a decline of motivation levels was observed between grade 7 and grade 9 for the students of the GUTS school as well as for the students in the national norm group.

Students at the GUTS school reported higher levels of well-being in grade 7 compared with students in the norm group. In contrast, well-being did not distinguish both groups in grades 8 and 9. A decline in well-being between grade 7 and grade 9 was visible for the GUTS cohort as well as for the national norm group.

For self-esteem a different pattern was detected: no differences between GUTS and the norm group were observed in grade 7, whereas the GUTS students reported somewhat lower self-esteem than the norm group in grades 8 and 9. Looking at the subscales of self-esteem more closely, it appeared that GUTS and the norm group did not differ on the two subscales of 'expression skills' (grade 7: $t(1577)=0.79$, $p=0.433$; grade 8: $t(1478)=1.71$, $p=0.087$; grade 9: $t(1048)=0.46$, $p=0.645$) and 'social skills' (grade 7: $t(1577)=1.94$, $p=0.052$; grade 8: $t(1478)=1.40$, $p=0.162$; grade 9: $t(1048)=0.63$, $p=0.530$). For the subscale 'confidence at tests', the GUTS group did not score differently in grade 7 ($t(1577)=1.32$, $p=0.188$), but did score lower than the norm group in grade 8 ($t(1478)=6.84$, $p<0.001$ and in grade 9 ($t(1048)=4.61$, $p<0.001$).

Discussion

This case study evaluated the effects of a learning environment (GUTS) on student performance and motivation in lower secondary education (grade 7 to 9). This learning environment specifically comprised a combination of intrinsic motivation stimulation and an

Table 5 Independent *t* test results for differences between GUTS and norm group for motivation, well-being and self-esteem in grades 7 to 9

Variable	Grade	GUTS		Norm group		Difference		
		M (SD)	*n*	*M* (SD)	*n*	*t*	*df*	*p*
Motivation	7	63.45 (7.48)	175	57.21 (10.10)	1404	7.91	1577	<0.001
	8	58.55 (8.23)	151	55.53 (9.93)	1329	3.60	1478	0.0003
	9	55.89 (9.99)	146	51.30 (11.31)	904	4.62	1048	<0.001
Well-being	7	66.29 (4.81)	175	62.53 (6.54)	1404	7.35	1577	<0.001
	8	60.65 (7.01)	151	61.02 (7.75)	1329	0.55	1478	0.5806
	9	60.49 (7.42)	146	59.86 (7.50)	904	0.96	1048	0.3388
Self-esteem	7	60.48 (7.84)	175	60.15 (7.98)	1404	0.53	1577	0.5992
	8	57.99 (8.46)	151	61.38 (9.29)	1329	4.29	1478	<0.001
	9	59.08 (9.02)	146	60.93 (8.26)	904	2.48	1048	0.0133

extrinsic incentive. The findings were remarkably consistent across the various comparisons and analyses. Overall, the GUTS cohort obtained higher report card grades than the comparison groups, either from the same school or from other schools. However, the performance decline observed at other schools and in previous cohorts of the same school was also observed in the GUTS cohort. The same combination of results was observed for student motivation: the GUTS cohort displayed higher motivation levels than the national norm group in all three grades. However, student motivation in the GUTS cohort declined between grade 7 and 9, and this decline is also visible in the data from the norm group. Collateral damage effects on students' sense of well-being and self-esteem were not observed, except for a specific effect on confidence on tests in eighth and ninth grade. These findings are discussed in more detail below.

Regarding student performance, the higher overall level of performance in the GUTS cohort shows that students generally are able to achieve at a higher level when we ask more of them, combined with special attention and support for the student's own preferences for particular school subjects. This finding supports the idea that the culture of C's is partially attributable to the fact that the school system does not require students to raise their performance. At the same time, performance levels declined between grades 7 and 9, as has been observed in a previous study (Wijsman et al. 2016). In the present study, the performance decline was particularly strong between grades 8 and 9. The report card grades were most positively affected in grades 7 and 8, whereas they dropped sharply between grade 8 and 9. The decline in this final year was stronger for GUTS students compared with students at other schools.

This particular pattern in the findings could be related to the following factors. First, it is possible that the curriculum in ninth grade becomes more demanding compared with previous grades. By ninth grade, students have learned the basics of school subjects and are expected to progress to higher-order learning tasks (Krathwohl 2002). These higher-order learning tasks require more complex skills from students, which can negatively affect results if students are unable to keep up. A second reason for declining performance in ninth grade, within Dutch education, is that students choose a certain combination of subjects (the so called 'profile choice') by the end of ninth grade for the remainder of their secondary school career. This means that every student can opt out of certain subjects after ninth grade if these subjects are not part of their profile choice. However, in ninth grade, all subjects still have to be taken. It is conceivable that student performance (and motivation) automatically declines in the ninth grade for discarded subjects. Third, in ninth grade, students are around the age of 15 years when puberty is at its peak. Pubertal status appears to have an indirect negative effect on achievement through academic motivation (Martin and Steinbeck 2017).

Regarding student motivation, GUTS students showed higher motivation levels throughout lower secondary education than students in the national norm group. This pattern mimics the result found for student performance and shows that it is possible to raise performance and motivation level at the same time. However, the decline of student motivation that has been found in previous studies (Eccles et al. 1991; Peetsma et al. 2005) was also observed in the GUTS cohort in the present study. Apparently, GUTS was not a strong enough intervention to counter the motivation decline. The autonomy-supportive elements included in GUTS might not have been sufficient to neutralise the mismatch between student needs and the school curriculum at large. The talent lessons were restricted to a few specific school subjects. In another study, we showed transfer of need-support in the talent lessons to the motivation in the corresponding school subject (Wijsman et al. 2018). However, this effect might not have generalised to motivation for school. In addition to

autonomy-support, which was provided in the talent lessons, fulfilment of the need for competence through guidance and feedback is also needed to stimulate motivation (Stroet et al. 2013).

For well-being and self-esteem, our study revealed no collateral damage to student well-being and self-esteem while trying to raise performance and motivation levels. Compared with a national norm group, GUTS students showed higher well-being levels at the start of their secondary education, whereas no difference was observed in grades 8 and 9. For self-esteem, it was observed that GUTS students were not different from the norm group on two of the subscales (expression and social skills) but, in grade 8 and 9, they scored lower on confidence at tests. This particular effect might be explained by the higher performance standard, which increases the expectations regarding students' performance. Asking more of students, on the one hand, could lead to increased perceived competence when one reaches this higher expectation and, on the other hand, could lead to low self-esteem or fear of failure when expectations are not within one's reach. In eighth and ninth grade, students might have felt less confident in reaching the higher performance expectation.

Regarding the underlying idea for GUTS, an explicit choice was made for an extrinsic incentive in terms of a higher performance norm. Taking the decreasing pattern of motivation and self-esteem for tests into account, it is possible that a different type of incentive could have led to other results, as the contingency of incentives is very important for its effects (Cerasoli et al. 2014; Deci et al. 1999). A type of incentive that is generally found to relate positively to intrinsic motivation is positive feedback (Deci et al. 1999; Hattie and Timperley 2007). A future iteration of GUTS could include more explicit attention to positive feedback when students show improving achievement on the way to meeting the higher performance standard.

Limitations

GUTS was implemented to decrease the effect of a culture of C's and to raise student performance and motivation through a combination of intrinsic and extrinsic motivators. A drawback of this approach is that we cannot ascertain the independent effects of both components on performance, motivation or both. The present study was a case study within a particular context. Future research could experimentally vary both components to ascertain their relative contribution with respect to the various outcome measures.

Between grade 7 and 9, 21 students (11.4%) dropped out from the sample because of insufficiently high grades. This percentage is not significantly higher than for previous non-GUTS cohorts at the same school. However, the dropout rate might have contributed to an overall positive effect of GUTS on performance levels.

The school was requested not to change the grading procedures, and the performance decline has shown that teachers did not inflate their grades to such an extent that the students' grades would not drop below a 7. Yet, we cannot be absolutely sure that individual teachers might have altered the norms used for grading tests, because we have not studied how the grading procedures were executed before and after the implementation of GUTS. Hence the possibility that unintended changes in grading procedures might have affected the results of this study cannot be excluded.

Implications and conclusions

Despite its limitations, the results of this case study provide a basis for teachers and school leaders to critically reflect on what motivates students in school to achieve. It appeared possible to increase achievement levels, which means that (some) students currently could be performing below their abilities. At the same time, performance and motivation declined over time, despite the school's great efforts to create a stimulating learning environment. GUTS might not have been the best way to achieve the intended results. On the other hand, we might also consider that this decline might be inevitable because of the hormonal effects of puberty, the changing needs in early adolescence, and the greater focus on peer relations. Further research is needed to disentangle developmental effects from the effects of a novel learning environment.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Cerasoli, C. P., Nicklin, J. M., & Ford, M. T. (2014). Intrinsic motivation and extrinsic incentives jointly predict performance: A 40-year meta-analysis. *Psychological Bulletin, 140*(4), 980–1008.

Danish Ministry of Education. (2014). *Improving the public school: Overview of reform of standards in the Danish public school (primary and lower secondary education).* Retrieved from Kopenhagen, Denmark: http://planipolis.iiep.unesco.org/sites/planipolis/files/ressources/denmark_improving_the_public_school.pdf. Accessed 2 May 2017.

De Boer, G. C., Minnaert, A. E. M. G., & Kamphof, G. (2013). Gifted education in the Netherlands. *Journal for the Education of the Gifted, 36*(1), 133–150.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125*(6), 627–668.

Dutch Ministry of Education Culture and Science. (2011). *Actieplan Basis voor Presteren; naar een ambitieuze leercultuur voor alle leerlingen [Action plan "Basis for performance; to an ambitious learning culture for all students"].* Retrieved from Den Haag, Netherlands: https://www.rijksoverheid.nl/documenten/kamerstukken/2011/05/23/actieplan-vo-beter-presteren. Accessed 20 Mar 2017.

Eccles, J. S., Lord, S., & Midgley, C. (1991). What are we doing to early adolescents—The impact of educational contexts on early adolescents. *American Journal of Education, 99*(4), 521–542.

Eccles, J. S., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development, 64*(3), 830–847.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112.

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). New York: Routledge.

Juvonen, J., & Murdock, T. B. (1995). Grade-Level Differences in the social value of effort: Implications for self-presentation tactics of early adolescents. *Child Development, 66*(6), 1694–1705.

Könings, K. D., Brand-Gruwel, S., & van Merriënboer, J. J. G. (2011). The match between students' lesson perceptions and preferences: Relations with student characteristics and the Importance of motivation. *Educational Research, 53,* 439–457.

Könings, K. D., Wiers, R. W., van de Wiel, M. W. J., & Schmidt, H. G. (2005). Problem-based learning as a valuable educational method for physically disabled teenagers? The discrepancy between theory and practice. *Journal of Developmental and Physical Disabilities, 17*(2), 107–117.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice, 41*(4), 212–218.

Lee, N. C., Krabbendam, L., Dekker, S., Boschloo, A., de Groot, R. H. M., & Jolles, J. (2012). Academic motivation mediates the influence of temporal discounting on academic achievement during adolescence. *Trends in Neuroscience and Education, 1*(1), 43–48.

Martin, A. J., & Steinbeck, K. (2017). The role of puberty in students' academic motivation and achievement. *Learning and Individual Differences, 53,* 37–46.
Mijs, J. J. B., & Paulle, B. (2016). The burden of acting wise: Sanctioned success and ambivalence about hard work at an elite school in the Netherlands. *Intercultural Education, 27*(1), 22–38.
Ministry of Education and Culture of Finland. (2014). *ET2020 National report of Finland*. Retrieved from Helsinki, Finland: http://ec.europa.eu/education/policy/strategic-framework/doc/et2020-national-report-fi_en.pdf. Accessed 2 May 2017.
OECD. (2016). *Netherlands 2016: Foundations for the future*. Paris: OECD Publishing.
Peetsma, T., Hascher, T., van der Veen, I., & Roede, E. (2005). Relations between adolescents' self-evaluations, time perspectives, motivation for school and their achievement in different countries and at different ages. *European Journal of Psychology of Education, 20*(3), 209–225.
Peetsma, T., & van der Veen, I. (2011). Relations between the development of future time perspective in three life domains, investment in learning, and academic achievement. *Learning and Instruction, 21*(3), 481–494.
Poorthuis, A. (2012). *Children in transition: Challenges and opportunities in adjusting to secondary school*. Utrecht: Utrecht University.
Reeve, J. (2006). Teachers as facilitators: What autonomy-supportive teachers do and why their students benefit. *The Elementary School Journal, 106*(3), 225–236.
Ryan, R. M., & Deci, E. L. (2002). An overview of self-determination theory: An organismic-dialectical perspective. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 3–36). Rochester: The University of Rochester Press.
Stroet, K., Opdenakker, M. C., & Minnaert, A. (2013). Effects of need supportive teaching on early adolescents' motivation and engagement: A review of the literature. *Educational Research Review, 9,* 65–87.
Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Boston: Pearson Education.
Tomlinson, C. A., Brighton, C., Hertberg, H., Callahan, C. M., Moon, T. R., Brimijoin, K., et al. (2003). Differentiating instruction in response to student readiness, interest, and learning profile in academically diverse classrooms: A review of literature. *Journal for the Education of the Gifted, 27*(2–3), 119–145.
Vansteenkiste, M., Sierens, E., Goossens, L., Soenens, B., Dochy, F., Mouratidis, A., et al. (2012). Identifying configurations of perceived teacher autonomy support and structure: Associations with self-regulated learning, motivation and problem behavior. *Learning and Instruction, 22*(6), 431–439.
Vorst, H. C. M., Smits, J. A. E., Oort, F. J., Stouthard, M. E. A., & David, S. A. (2010). *SchoolVragenLijst voor basisonderwijs en voortgezet onderwijs; handleiding en verantwoording*. Amsterdam: Pearson Assessment and Information B.V.
Warrington, M., Younger, M., & Williams, J. (2000). Student attitudes, image and the gender gap. *British Educational Research Journal, 26*(3), 393–407.
Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68–81.
Wijsman, L., Warrens, M. J., Saab, N., van Driel, J. H., & Westenberg, P. M. (2016). Declining trends in student performance in lower secondary education. *European Journal of Psychology of Education, 31,* 595–612. https://doi.org/10.1007/s10212-015-0277-2.
Wijsman, L., Saab, N., Schuitema, J. A., van Driel, J. H., & Westenberg, P. M. (2018). Student perceptions of autonomy support and structure in enrichment classes: Relationship with subject motivation and self-efficacy (submitted).
Yin, R. K. (2011). *Applications of case study research*. Thousand Oaks, CA: Sage.