The role of students' interests during computer-assisted learning: A field experiment

Computers in Human Behavior 130 (2022) 107168

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

The role of students' interests during computer-assisted learning: A field experiment

Kaat Iterbeke [a,*], Wouter Schelfhout [b], Kristof De Witte [a,c]

[a] Leuven Economics of Education Research, KU Leuven, Belgium
[b] Antwerp School of Education, Universiteit Antwerpen, Belgium
[c] UNU-MERIT, Maastricht University, the Netherlands

ARTICLE INFO

Keywords:
Computer-assisted learning
Interest
Example choice
Personalization
Financial education

ABSTRACT

The present paper examined the effect of adjusting learning material to the out-of-school interests of students. In particular, using a field experiment involving 1449 secondary school students in 31 Flemish schools, we evaluated the effectiveness of the instructional method "example choice" in a computer-assisted financial education program. Example choice allowed students to choose between contexts potentially more aligned with their out-of-school interest, which in turn determined the backdrop examples for exercises in the computer-assisted program. Schools were randomly assigned to a control or treatment condition. While students in both the control and treatment schools received the computer-assisted program, example choice was added only to the program for students in treatment schools. Results show that example choice did not increase students' financial knowledge. Despite the well-established psychological belief that students' interest in an academic topic can be instilled by the educator or the learning environment, example choice did not affect the interest of the average student in the financial topic. Moreover, example choice even led to significant motivational deficits for students with low perceived competence.

1. Introduction

Despite the high private returns to investment in education, many students devote only low levels of effort to their schooling. As an explanation for this puzzle, an increasing body of evidence indicates that non-cognitive skills play an equally important role for education investments as cognitive skills (Koch et al., 2015). Self-confidence, for instance, is suggested to play a key role in building up intrinsic motivation such as curiosity and a joy of learning. Similarly, students' interest in the academic topic can increase effort and productive learning behavior (Bernacki & Walkington, 2018). An important question for the latter, however, is how to elicit interest in a particular academic topic and how computer-assisted learning might facilitate this.

As the out-of-school interests of students differ within a single classroom, psychological research has suggested these interests may serve as effective tools to support learning. Not only does customizing features of the material to students' out-of-school interests allow the curriculum to become more meaningful to students (Subban, 2006), these interests are also linked to students' motivation and learning in the short and long term (Tomlinson et al., 2003). Multiple interest-based interventions exist, one of which is example choice. Example choice is an instructional method that, for a given academic topic, offers students a choice between predefined contexts related to their out-of-school interests (Reber et al., 2009). Given the opportunity technology offers to customize learning material, the effectiveness of example choice has been evaluated in computer-based learning environments. In particular, technology can support teachers by easing the process of assessment of students' interests, by providing instructional suggestions to students, and by delivering engaging (often game-based) interactive content (Muralidharan et al., 2019).

Several experimental studies examined example choice in an educational setting. The majority of the existing studies reported a significant increase in students' interest in an academic topic, though no increase in learning (Høgheim & Reber, 2015, 2017; Reber et al., 2009). One exception is a recent study by Clinton and Walkington (2019) who reported significant positive effects of example choice on students' performances. It should be noted, however, that, apart from the latter study, previous experiments were conducted in a laboratory setting (i.e., students completed the program in computer labs, with experimenters present during the session). Hence, little is known about the effects on students' performance and interest in an academic topic in a classroom environment. Moreover, given all studies were based on a single learning session in which students

* Corresponding author. Leuven Economics of Education Research, KU Leuven, Naamsestraat 69, 3000, Leuven, Belgium.
E-mail address: Kristof.dewitte@kuleuven.be (K. Iterbeke).

https://doi.org/10.1016/j.chb.2021.107168
Received 18 June 2021; Received in revised form 7 November 2021; Accepted 21 December 2021
Available online 27 December 2021

were offered a single choice, the extent to which example choice can be used in the learning material remains largely unexplored.

Using a field experiment in the seventh and eighth grades of Flemish secondary education involving 1449 students in 31 schools, we evaluate the effect of example choice during multiple learning sessions including multiple choice moments in a classroom environment. By offering students the choice between different contexts multiple times throughout the learning material, our approach differs from previous experimental studies (Clinton & Walkington, 2019; Høgheim & Reber, 2015; 2017; Reber et al., 2009) that utilized a shallow form of customization in which the context picked by students was held constant during the entire learning session. The computer-assisted learning material focussed on financial education, in particular, the role of the government and taxes in financial systems. This is an interesting topic for our study as interest-based interventions are designed to provide a type of grounding, making complex topics more situated and understandable (Walkington, 2013). Given that students in the seventh and eighth grades are assumed to have little interest and knowledge about the role of government and taxes, relating this abstract topic to concrete objects or events that students are familiar with, such as their out-of-school interests, might benefit them, both in terms of enhancing their financial knowledge and increasing their (re-)engagement with the financial topic. This notion is also supported by research on financial education. However, while the use of more entertaining learning material, relevant to the lives of students, is supported by literature on financial education, empirical evidence on the effects is lacking (Kaiser & Menkhoff, 2020).

Accordingly, the contribution of this study is twofold. First, we contribute to the literature by examining the extent to which example choice can be used in "real life" settings, i.e., offering students multiple choice moments during multiple learning sessions in a regular classroom environment. More realistic conditions are particularly valuable for policymaking. Second, the study contributes to the financial education literature by exploring the importance of drawing on students' out-of-school interests when teaching financial topics.

The paper is organized as follows. First, we provide a theoretical framework for interest and example choice. Next, we describe the experiment. Third, we present the empirical strategy and main results, before offering concluding remarks in the final section.

2. Theoretical framework

2.1. Interest theory

Theorists have defined two types of interests, i.e., individual and situational interest. *Individual interest* is an enduring state of preference for a certain topic. It can be emerging or well-developed (Hidi & Renninger, 2006). *Situational interest* is an attention-heightened and affective reaction to particular features of an environment that the student is personally connected with (Walkington, 2013). According to Hidi and Renninger (2006), situational interest mirrors the two first phases of a four-phased model, which can evolve into an enduring individual interest. That is, situational interest can be stimulated by features of an environment – *triggered situational interest* – and maintained over time as the student engages further with the stimuli. The latter can be differentiated into two dimensions, i.e., an affective-based connection to the content – *maintained situational interest* and a value-based connection to the content – *perceived value* (Linnenbrink-Garcia et al., 2010). In an academic context, the triggered situational interest refers to students' opinion of the form of the instruction, whereas maintained situational interest refers to students' opinion of the content of the instruction. The theory of interest thus suggests that an educator may elicit situational interest in an academic topic by changing factors in the learning environment. For instance, an educator can adjust features of the learning environment by using the out-of-school interests of students (i.e., example choice) to trigger students' situational interest in the academic topic, which might be maintained and eventually evolve into an individual interest in the academic topic. In Fig. 1, we show the interlinkage between these different types of interest.

2.2. Example choice

Example choice can be defined as an instructional method that offers students choices between predefined contexts. Students can select contexts that they find most interesting, which in turn determine the backdrop examples used in the exercises they are working on (Reber et al., 2009). For instance, in the context of financial education, an exercise about government expenses can be tied to students' out-of-school interests by providing context-related rather than standard examples of government expenses (e.g., sports-related if that is what the student finds most interesting).

On the one hand, example choice may result in positive learning outcomes. In Reber et al. (2009), three potential mechanisms were outlined through which example choice can improve student outcomes. The mechanisms are not mutually exclusive. First, if students select a context they are interested in, and this context increases interest in the academic topic, students' persistence in learning will improve. Personalized contexts are thought to trigger students' situational interest in the academic topic. This, in turn, can become maintained as students enjoy the personal connection or the connection to the out-of-school interests they value – the two dimensions of maintained situational interest (Bernacki & Walkington, 2018). Second, if students select a context they know the most about, prior knowledge may affect learning. Personalized contexts can activate students' prior knowledge by grounding the academic topic in a context that is familiar to the student. As a result, access to long-term memory may be facilitated due to familiarity; students are less likely to make conceptual errors as their out-of-school knowledge provides insights into the reasonableness of answers; and concrete details can make the learning material more understandable (Bernacki & Walkington, 2018). Finally, example choice may increase students' perceived control. Allowing students to make choices will increase their sense of autonomy and weaken their perception of teacher control (Flowerday et al., 2004) and in turn may positively affect intrinsic motivation (Ryan & Deci, 200), situational interest in the academic topic (Linnenbrink-garcia et al., 2013), and learning (Cordova & Lepper, 1996; Schneider et al., 2018).

On the other hand, there are several reasons to expect that example choice has no or negative effects on educational outcomes. First, example choice may operate as *seductive details*, i.e., elements that students may perceive as highly interesting, but which are distracting them from what needs to be learned. Research has indicated that seductive details, due to redundant cognitive processing demands, can have a detrimental effect on performance (Harp & Mayer, 1998; Lehman et al., 2007), in particular among the weaker students (Magner et al., 2014). This idea is supported by the Cognitive Load Theory (Sweller et al., 1998), which states that additional information (which is irrelevant for the learning goal) can overload the working memory of students. Second, Schwartz (2000) argued that offering multiple choices can be unattractive due to judgments that involve information that is inaccessible or unavailable for the student. In particular, this may lead to a cognitive burden (Patall et al., 2008), thereby decreasing students' sense of personal autonomy and motivation (de Brabander & Martens, 2014). In a similar vein, Katz and Assor (2007) explained choice effects based on the Self-Determination Theory (Ryan & Deci, 2000) and argued that the choices must be *competence-enhancing* (i.e., not too numerous or complex) to be effective. Concerning example choice, however, it is unclear at which point choices become too numerous or complex.

To explain differences in the effectiveness of interest-based interventions, Clinton and Walkington (2019) argued that manipulating the *depth* (whether the intervention is designed to focus on deep knowledge or surface features of students' out-of-school interest), *grain size* (whether the intervention is designed to be broadly or individually customized), and *ownership* (whether students take an active role in personalizing the intervention) of interventions appears to significantly

influence outcomes. Given that our study differs to some extent from previous research in terms of these three factors, and based on the potential mechanisms outlined above, the effects of example choice, as designed in our study, are *a priori* ambiguous.

Finally, both perceived competence and individual interest in the academic topic appear to moderate the effect of choice in learning activities, though no univocal evidence is found. For instance, Patall (2013) and Patall et al. (2014) revealed beneficial effects of choice for students with high levels of perceived competence and individual interest in the academic topic, while Høgheim and Reber (2015) found positive effects for students with low individual interest in the academic topic. Accordingly, we investigate the importance of both factors in our study.

3. Materials and methods

3.1. Computer-assisted learning material

The financial education program was designed as three lectures of 50 min in the form of an interactive website. Students completed the material individually. To prevent additional teacher influences on the program effectiveness, we minimized the role of the teacher by using automatic differentiated instruction and feedback on the website. However, in contrast to previous experiments (Høgheim & Reber, 2015, 2017; Reber et al., 2009), the teacher remained present in the classroom to maintain a natural classroom environment for students. If necessary, the teacher provided guidance to students or prompted them to keep on working.

The learning content in the computer-assisted learning material focussed on the topic of the government and taxes, which is part of the content area of the PISA financial literacy assessment, *Planning and Managing Finances* (OECD, 2016). After the program, students were expected, among others, to understand how government taxes and benefits affect personal finances, and identify various types of income (such as allowances and salary) and ways of discussing income (gross and net income). The website consisted of three chapters and each chapter included seven exercises, multiple information sheets, and a formative test. The latter was included to inform students about the learning goals of each chapter and the extent to which they had reached these goals. Formative tests are suggested to stimulate learning without serving as an actual qualifying instrument (Ghysels et al., 2014). The exercises comprised, among others, multiple-choice quizzes, fill-in-the-blank and drag-and-drop exercises, learning games, interactive videos, and case studies from which students learned from explanations and feedback. Students had to repeat an exercise until they answered correctly.

3.2. Intervention

We designed two conditions to study whether modifying material to draw on students' out-of-school interests improves students' financial knowledge and interest in the financial topic. In the control condition, students followed the computer-assisted program including generic exercises without any reference to out-of-school interests. In the treatment condition, students followed the computer-assisted program in which, for multiple exercises, they were given the choice between different contexts based on three (out-of-school) interest fields, i.e., sports, music and culture, and social media and gaming.[1] Two senior teachers in the research team who have ample experience in triggering the interest of students developed the learning material. To link the learning material in the treatment condition with the out-of-school interests of students, the three interest fields were chosen based on the responses on a short survey that was completed by a random sample of almost 250 students in the seventh and eighth grades in different Flemish secondary schools (see Table B.1 in Appendix for the descriptive statistics of the survey). Contrary to earlier literature (e.g., Reber et al., 2009), the learning material in the treatment condition did not aim to connect the learning content with elements of students' everyday life such as foods or friends, but with students' interests in a specific topic (as in Clinton & Walkington, 2019; Høgheim & Reber, 2015; 2017). Moreover, in contrast to the previous experimental studies in which the context picked by students was held constant during the learning session, the degree of choice was maximized in our setting, i.e., students were offered seven choice moments in the learning material. Though the delivery approach was modified, the learning objectives, core content, or difficulty level of the program were not altered. The senior teachers in the research team tried to match the linguistic structure of both versions of the material as well as possible. Figure A.1 and Figure A.2 in Appendix provide examples of the differences in the learning material.

Teachers were sent the learning material by e-mail and had to plan the three lectures during regular class hours within six weeks after receiving the learning material. Accordingly, the time between lectures could differ across classes and schools. Students had to take a pre-and post-treatment test. A schematic overview of the difference in the conditions is provided in Fig. 2.

3.3. Test instruments

We test the effectiveness of the program via two computer-assisted tests from which we derived multiple outcome measures, i.e., financial knowledge, triggered situational interest, maintained situational interest, and perceived value of the financial topic.

First, to capture students' baseline knowledge and attitudes about the financial topic, students completed a pre-treatment test before the start of the lectures. The test included several items concerning students' demographics such as their age, gender, and native language. To measure initial differences in students' individual interest in the financial topic, we used four items, i.e., "*I think the topic of government and taxes is meaningless*" (reversed scored), "*Given the topic of government and taxes is fun, I would like to know more about it*", "*I think the topic of government and taxes is boring*" (reversed scored), and "*Learning about the government and taxes is important to me*". This measure was inspired by Høgheim and Reber (2015). Cronbach's alpha was 0.76. Students' perceived competence in the financial topic was measured via the item "*I have sufficient knowledge about the government and taxes*". Responses to all items were given on a five-point Likert scale ranging from "totally disagree" (1) to "totally agree" (5). In addition, to assess students' initial financial knowledge, the pre-treatment test consisted of eleven multiple-choice questions that referred directly to the material (similar to questions 1–11 in Appendix C). Four other questions, related to rates of return, (compound) interest, and inflation, were included to measure students' financial literacy more generally (similar to questions 12–15 in Appendix C). These latter questions were taken from Lusardi and Mitchell (2011). Cronbach's alpha for the pre-test knowledge questions was 0.65. The mediocre test reliability can be explained by the limited amount of items assessing a broad range of knowledge. Except for two questions showing poor discrimination, all questions had adequate discrimination indices (Pearson Product Moment correlation of 0.22 or higher). The difficulty indices ranged from 0.08 to 0.68.

Second, to measure the effectiveness of the different versions of the program, a post-treatment test was administered shortly after the

[1] In total, seven out of 21 exercises were adjusted to students' out-of-school interests. The adjustments to students' out-of-school interests were made depending on the type of exercise. For instance, the games and videos were identical in both versions of the program. While the differences between both versions seem to cover only one-third of the program, it should be noted that, in general, the interest-adjusted exercises for students in treatment schools and corresponding generic exercises for students in control schools took the most time to complete. Accordingly, we believe that there was sufficient variation in the conditions.

lectures.[2] The knowledge questions in the test measured similar concepts and were created by rephrasing and using adjustment of numbers of the pre-test questions. Further, we measured students' situational interest in the financial topic using three scales, i.e., triggered situational interest, maintained situational interest, and perceived value, inspired by the *Student Interest Survey* (Høgheim & Reber, 2015; Linnenbrink-Garcia et al., 2010). Appendix C provides a detailed description of the underlying items of each scale. The triggered situational interest scale aimed to capture students' opinions of the learning material. Maintained situation interest referred to students' affective experiences related to the learning material. Note that maintained situational interest differs from triggered situational interest because students' enjoyment of working with the learning material is based on the content (i.e., the financial topic) rather than the specific features of the learning material (Linnenbrink-Garcia et al., 2010). The perceived value scale measured students' value perception of the financial topic. Cronbach's alpha was 0.88 for triggered situational interest, 0.86 for maintained situational interest, and 0.71 for perceived value. We conducted a Confirmatory Factor Analysis (CFA) using Structural Equation Modeling (SEM) to verify the factor structures of the situational interest scales. The overall measurement model provided a good fit based on the RMSEA, CFI, TLI, and SRMR criteria (RMSEA = 0.089, CFI = 0.932, TLI = 0.912, SRMR = 0.062).

Finally, we gauge the fidelity of implementation via eight classroom observations conducted by our research team and several evaluation surveys completed by teachers after the lectures. When allowed by school regulations, the classroom observations were filmed.

3.4. Sample

The computer-assisted program was developed for students in the seventh and eighth grades of Flemish secondary education. Schools were recruited via an open call in December 2018. Schools were given the choice to participate in one of two waves, i.e., a first wave from mid-February 2019 to end-March 2019 and a second wave from mid-April 2019 to end-May 2019. By offering the choice between the two waves, we expected to attract a larger number of schools. Fig. 3 provides a timeline of the study.

Our sample comprised 1449 students in 31 schools who completed the pre-treatment test. The 31 schools were located across all the different provinces of Flanders, however, most schools were located in the provinces Antwerp (7 schools) and East Flanders (14 schools). 21 out of 31 schools were located in larger cities. To assess the external validity of our sample of schools, we compared the school characteristics of in-sample and out-of-sample schools using administrative data on Flemish secondary schools. The first panel of Table B2 in Appendix shows that the 31 schools were no different than the average Flemish school with respect to four commonly used socioeconomic indicators, i.e., the percentage of children with a mother without a secondary education degree, the percentage of children not speaking Dutch (the official language) at home, the percentage of children receiving an allowance, and the percentage of children living in a neighborhood with high retention rates (defined as students whose study falls two years behind schedule at the age of fifteen).[3] Hence, our baseline sample of schools was representative of the average Flemish school.

3.5. Randomization

The 31 schools were randomly assigned to one of the two conditions.[4] Randomization was done at the school level to implement the computer-assisted program in students' natural classroom environment without the threat of contamination between teachers in different conditions. The control condition consisted of 727 students and the treatment condition of 722 students, such that the sample size provided sufficient power. While randomization was stratified by school size to ensure balanced sample sizes, schools in control and treatment condition were also comparable with respect to the type of education (private/public) and the share of participating eighth-grade students in the school, as presented in panel A of Table 1.[5] Apart from the individual interest of students in the financial topic, student characteristics were similar. Students in the control condition reported having a higher interest in the financial topic. Most importantly, however, the baseline test score was not significantly different and amounted to 6.36 out of 15 on average.

3.5.1. Attrition

While 1449 students completed the pre-treatment test, only 667 (46 percent) also took the post-treatment test. The follow-up rate was 49 percent in the control condition and 43 percent in the treatment condition. As presented in panel A of Table 1, the rate of attrition was not significantly different across the two conditions. When looking at the percentages of compliance within classes, we found that 79 percent of the attrition can be explained by teachers not following the prescribed instructions, i.e., in 42 out of 90 classes, no student completed the test. Given this high attrition rate, we explored non-compliance more in-depth using the data on the fidelity of implementation. The classroom observations and teacher evaluation surveys revealed no deviations from the prescribed instructions that would raise validity issues for our empirical analysis. However, several teachers reported that the material was challenging for the studied age group. Consequently, students might not have finished the material in time to complete the post-test. Alternatively, though a detailed lesson guide was provided to the teachers, the classroom observations revealed that the formative tests after each chapter of the learning material might have been mistaken for the post-treatment test. These findings suggest that the non-compliance of teachers can be explained by the design of the learning material.[6]

To assess the validity of randomization and the importance of attrition for the following regression analyses, we regressed each baseline characteristic on a constant, a treatment indicator, an attrition indicator, and an interaction term of the latter two (as in Lai et al., 2015). Table B.4 shows that non-complying students (who did not complete the post-test) were older, had a significantly lower socioeconomic status (SES), and had higher perceived competence in the financial topic than complying students (who completed both the pre- and post-test). However, given the estimates of the interaction terms are insignificant, we do not believe that attrition will bias our main results as there was no selective attrition in the two conditions.

In panel B of Table 1, the descriptive statistics of the final sample are

[2] We asked teachers to plan the post-treatment test after the third lecture and within the six weeks after receiving the learning material. Although we do not have exact information about the time between the third lecture and the post-treatment test, we did ask teachers during the registration how they were planning to implement the lectures, i.e., in one block, two blocks, or four separate blocks of lecture hours. An analysis of this data suggested no significant differences between the control and treatment conditions (see Table B.3).

[3] AGODI, Cijfermateriaal - Leerlingenkenmerken (2018–2019), available at http://www.agodi.be/cijfermateriaal-leerlingenkenmerken.

[4] Note that the randomization of schools was conditioned upon completion of the pre-treatment test by students in the school.

[5] The type of education refers to the different educational networks in the Flemish education system. There are three networks, i.e., (1) public official education organized by the Flemish Community, (2) public government-aided education managed by municipal or provincial authorities, and (3) government-aided education managed by a private person or organization (largest network). In the remainder of the paper, we consider the two former categories as 'public education providers', and the latter as a 'private education provider'.

[6] To formally test this assumption, we check whether the non-compliance of teachers was random by comparing school-, class-, and teaching characteristics of complying and non-complying teachers across conditions. Table B.3 shows that this source of attrition is random.

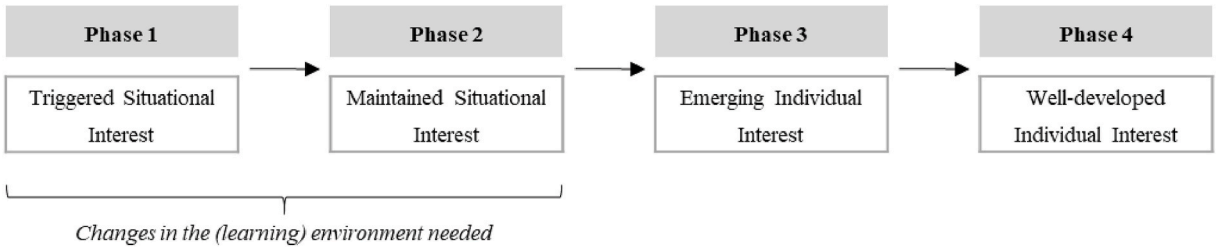


Fig. 1. Interest development.

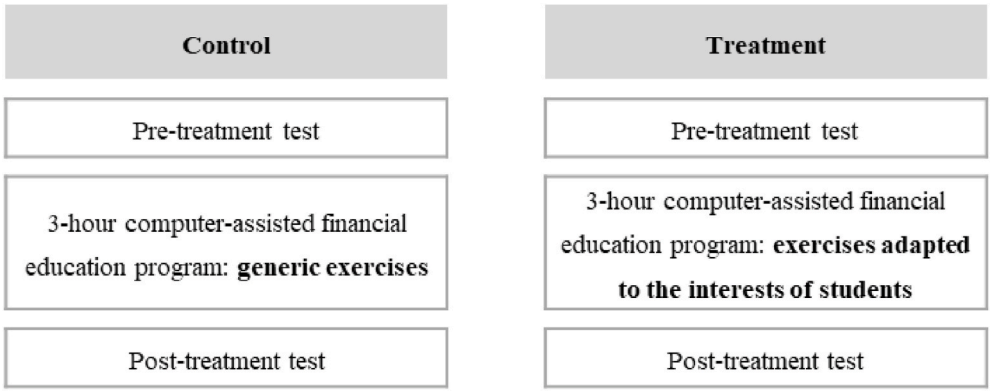


Fig. 2. Experimental conditions.

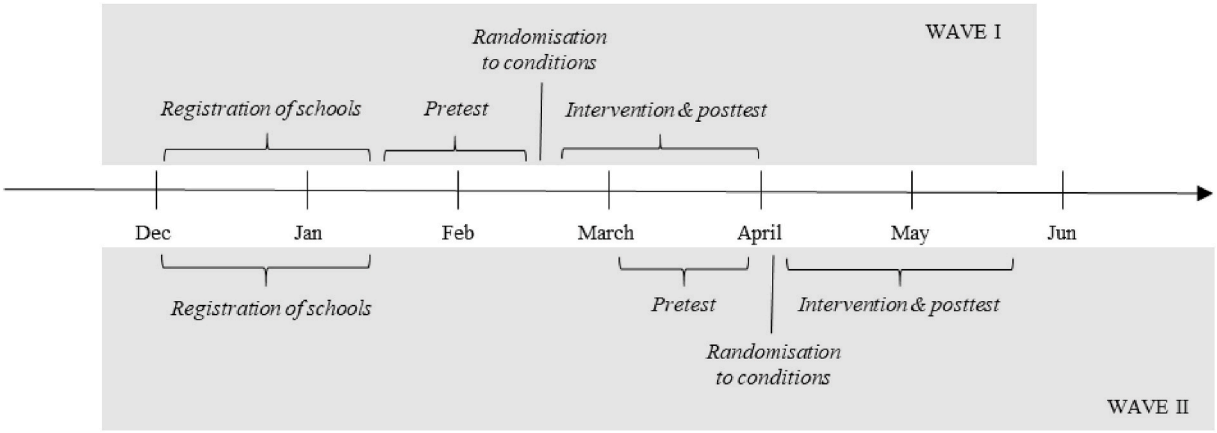


Fig. 3. Study timeline.

presented. The panel shows that the conditions included students with similar background characteristics. Most students were in the eighth grade (97 percent) and followed an academic track (93 percent). 46 percent of the students in the final sample were female and 85 percent spoke Dutch at home. While students had some prior interest in the financial topic (mean value of 3.26 out of 5), their perceived competence was 2.7 out of 5, on average. The baseline test score was 6.5 out of 15, on average. Finally, the follow-up test scores, triggered situational interest, maintained situational interest, and perceived value offered us a first indication of the effectiveness of the computer-assisted program adjusted using example choice as compared to the computer-assisted program without example choice. Surprisingly, students receiving the computer-assisted program without example choice reported having a higher triggered situational interest during the lectures, though the estimate is significant only at the ten percent level.

4. Results

4.1. Intent-to-treat effects

Before analyzing the effect of example choice, we compare the pre- and post-test scores of students in the treatment and control conditions to evaluate whether they learned from the material. Our research design did not include a control condition where students did not follow any program (i.e., both conditions received the learning material) because of three reasons. First, adding a third condition would reduce the sample size of the treatment and control cells. Second, from earlier experimental evidence on financial education, we know that there are no learning effects between the pre-and post-test for students receiving no financial education (Iterbeke et al., 2020).[7] Third, we are interested in the effect of example choice rather than the effect of the learning material itself. With these motivations in mind, Fig. 4 shows the

[7] Note that we designed the pre- and post-test in such a way that it minimized the learning effects from the tests.

Table 1
Sample descriptives.

	Control		Treatment		Difference in means test (*p*-value)
	Mean	SD	Mean	SD	
Panel A. All students at baseline					
School characteristics					
Private education	0.84		0.80		0.717
Fraction of 8th grade students	0.97		0.95		0.607
Background characteristics					
Track (Academic)	0.95		0.90		0.549
Age	13.10	0.42	13.12	0.50	0.713
Gender (female)	0.48		0.49		0.847
Language (Dutch)	0.79		0.86		0.286
SES (4)	3.00	1.03	3.02	1.02	0.879
Individual interest (5)	3.34	0.78	3.17	0.78	0.064
Perceived competence in topic (5)	2.75	0.95	2.71	0.93	0.635
Baseline test score (15)	6.39	2.83	6.33	2.94	0.912
Government and taxes (11)	4.67	2.21	4.53	2.28	0.703
Financial literacy (4)	1.72	1.06	1.80	1.12	0.618
Assigned number of schools	15		16		
Assigned number of students	727		722		
Attrition	0.51		0.57		0.672
Panel B. Students at follow-up					
School characteristics					
Private education	0.91		0.70		0.255
Fraction of 8th grade students	0.95		1.00		0.367
Background characteristics					
Track (Academic)	0.95		0.91		0.717
Age	13.03	0.39	13.09	0.42	0.419
Gender (female)	0.46		0.47		0.850
Language (Dutch)	0.82		0.89		0.454
SES (4)	3.14	1.00	3.15	0.96	0.955
Individual interest (5)	3.33	0.76	3.19	0.80	0.285
Perceived competence in topic (5)	2.70	0.95	2.71	0.90	0.933
Baseline test score (15)	6.47	2.92	6.56	2.91	0.899
Government and taxes (11)	4.70	2.28	4.66	2.27	0.942
Financial literacy (4)	1.77	1.07	1.89	1.08	0.484
Post-test test score (15)	7.93	3.03	7.54	3.10	0.684
Government and taxes (11)	5.87	2.41	5.51	2.38	0.620
Financial literacy (4)	2.06	1.09	2.03	1.19	0.915
Triggered situational interest (5)	3.59	0.84	3.28	0.90	0.085
Maintained situational interest (5)	3.15	0.80	3.00	0.82	0.186
Perceived value (5)	3.51	0.69	3.39	0.67	0.278
Assigned number of schools	10		9		
Assigned number of students	358		309		

Note. Scores are reported before standardisation; Value in parentheses after a characteristic refers to maximum value; Values for track refer to pre-vocational education (0) and academic education (1); SES is approximated by the number of times a student travels abroad during the year (1 = not, 2 = 1 time, 3 = 2 times, 4 = more than 2 times); *p*-value difference in means is derived from a regression of the characteristic on a treatment dummy with standard errors clustered at school level; 22 missing values for triggered situational interest; 21 missing values for maintained situational interest and perceived value.

effectiveness of the computer-assisted financial education program for students' financial knowledge in treatment and control schools.[8] We find that the overall test score (including both questions that related to the government and taxes and general financial literacy) and the subscore on the government and taxes improved for the two conditions after the program. In Table B.5 in Appendix, we formally test this finding via a Difference-in-Differences estimation and find that the scores increased significantly by half of a standard deviation on average. The improvement in test scores was lower for students in the treatment condition, though this observed difference is not significant.

Next, we estimate the effect of example choice using the following regression model,

$$y_{is}^{1} = \alpha + \beta_0 Treatment_{is} + \beta_1 y_{is}^{0} + \Sigma \beta_2^{'} X_{is} + \varepsilon_{is} \quad (1)$$

where $Treatment_{is}$ is an indicator for the assignment to the treatment condition (example choice) for student i in school s. y_{is}^{1} is the standardized value of an outcome measure, i.e., post-test test score, triggered situational interest, maintained situational interest, and perceived value, for student i in school s. y_{is}^{0} denotes the standardized baseline test score for student i in school s. X_{is} refers to the standardized values of the individual interest and perceived competence in the financial topic for student i in school s. Given randomization occurred at the level of the school, is clustered at school level s to allow for within-cluster dependence.

The results in Table 2 show that providing example choice in financial education did not enhance the learning experience of the average student. The overall test score and score related to the government and taxes of the treatment condition are 0.15 standard deviations lower than those of the control condition, though these estimates are not significant. Students' situational interest in the financial topic also remained unaffected when students were offered example choice in the learning material, i.e., the treatment reduces students' triggered situational interest, maintained situational interest, and perceived value by 0.33, 0.11, and 0.10 standard deviations, respectively, though these estimates are not significant. Despite the zero effect of example choice, it is worth noting that students' individual interest in the financial topic appears to be a good predictor of students' situational interest in the financial topic.

To assess the robustness of our main results, we perform four analyses. First, Table B.6 in Appendix shows that the estimates remain largely robust against the inclusion of additional control variables. Second, although the observed attrition rates are unlikely to result in biased estimates (as discussed in section 3.6), we perform a Lee (2009) bounds analysis. The estimates related to the test scores are robust, as presented in Table B.7 in Appendix. For the three situational interest measures, however, the lower bound estimates are significantly negative. Consequently, these findings must be interpreted with caution. Third, as an alternative check, we use the coarsened exact matching (CEM) technique proposed by Iacus et al. (2008) and Blackwell et al. (2009). CEM reduces the imbalance between the treatment and control conditions by matching students in both conditions using temporarily coarsened data. The results in Table B.8 in Appendix show that the estimates from the main analysis are similar to the estimates using matched samples, suggesting no confounding influences of covariates. Finally, as the inference with clustered standard errors assumes the number of clusters to go to infinity, the clustered standard errors reported in Table 2 might not be reliable due to the few clusters in our data (i.e., 19 schools). Hence, we account for this using the wild cluster restricted bootstrap approach proposed by Cameron et al. (2008). In particular, the approach applies bootstrapping to obtain critical values that provide an asymptotic adjustment for the few clusters. Table B.9 shows that the unadjusted and adjusted critical values are comparable.

4.2. Heterogeneous effects

Previous literature has indicated that students' individual interest and perceived competence in an academic topic moderate how students

[8] Since we do not have baseline values for the outcome measures related to students' situational interest in the financial topic, this analysis is restricted to students' test scores.

Table 2
Effects on students' academic performance and situational interest.

Dependent variable	Overall score	Score G&T	Triggered SI	Maintained SI	Perceived value
Treatment	−0.151	−0.145	−0.333	−0.110	−0.101
	(0.247)	(0.252)	(0.198)	(0.106)	(0.124)
Baseline test score	0.583***	0.519***	0.136***	0.156***	0.136***
	(0.0554)	(0.0501)	(0.0370)	(0.0295)	(0.0454)
Individual interest	−0.0134	0.0217	0.139***	0.404***	0.343***
	(0.0349)	(0.0463)	(0.0477)	(0.0346)	(0.0510)
Perceived competence	−0.0823**	−0.0657**	0.0797	0.0725**	0.000326
	(0.0326)	(0.0309)	(0.0651)	(0.0331)	(0.0381)
Observations	667	667	645	646	646
R^2	0.303	0.248	0.087	0.227	0.153

Note. Abbreviations G&T and SI refer to the questions related to the government and taxes and situational interest, respectively; Clustered standard errors at school level in parentheses; ***p < 0.01, **p < 0.05, *p < 0.1; Baseline test score in column 2–5 refers to the G&T score.

experience example choice and the effect of choice in general (Høgheim & Reber, 2015; Patall, 2013; Patall et al., 2014). Accordingly, we examine potential treatment heterogeneity by these moderators via interactions with the treatment, as presented by Tables B.10 and B.11 in Appendix. On the one hand, we find that the treatment effect was not affected by students' individual interest in the financial topic, i.e., the interaction term between the treatment and individual interest is insignificant for all five outcome measures. On the other hand, students' perceived competence in the financial topic appears to have moderated the impact of example choice on students' triggered and maintained situational interest. To further explore the nature of the significant interactions, we conduct simple effect tests (post-estimation tests that examine the effect of the treatment at specific levels of the moderator), as presented in Fig. 5. While high values of perceived competence did not moderate the impact of example choice, the instructional method appeared to significantly reduce the triggered and maintained situational interest of students with low perceived competence by 0.44 and 0.20 standard deviations, respectively.

Given the zero (or in some cases, negative) effect of example choice, it is important to consider this instructional method as *seductive details,* i.e., elements that students may have perceived as highly interesting, but which were distracting them from the core content of the learning material. As research has indicated that seductive details can have a detrimental effect on performance, in particular among the weaker students (Magner et al., 2014), Table B.12 in Appendix shows the interaction between the treatment and baseline test scores. We observe no statistically significant interaction for all five outcome measures.

5. Discussion

The present paper examines the effect of example choice in a computer-assisted financial education program using field experiment data involving 1449 students in 31 Flemish schools. Example choice enabled students to contextualize the learning material to three popular interest fields. We find that the post-test scores of students were significantly higher than their pre-test scores regardless of the instructional method used (i.e., for both the treatment and control condition). Further, example choice did not affect students' situational interest in the financial topic on average and led to significant motivational deficits for students with low perceived competence in the financial topic. The latter finding is in line with Patall et al. (2014) who argue that, without some initial sense of expertise, greater freedom of choice may lead to motivational deficits.

Our results are partially consistent with the existing evidence on example choice. Except for Walkington and Clinton (2019), no gains in students' knowledge were reported in previous studies (Høgheim & Reber, 2015; 2017; Reber et al., 2009). In contrast to our study, students' interest in the academic topic significantly increased in all previous studies. The subject (i.e., financial education) and the length of our intervention differed from the previous studies, which all investigated

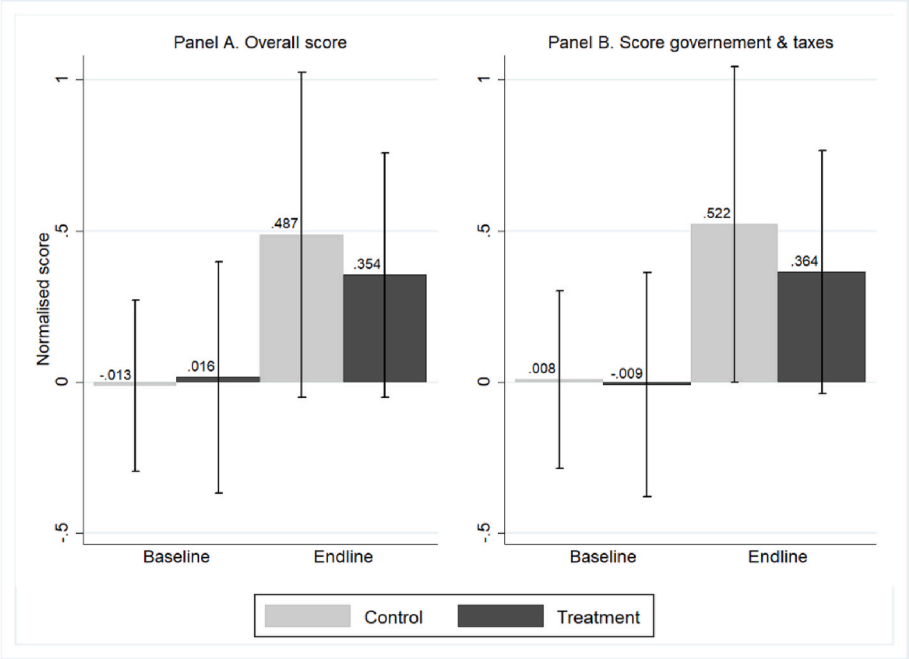


Fig. 4. Mean Test Scores in Treatment and Control Condition. *Note.* The test scores, normalised with respect to baseline, are presented across the treatment and control condition with the 95 percent confidence intervals, adjusted for within-cluster dependence at school level.

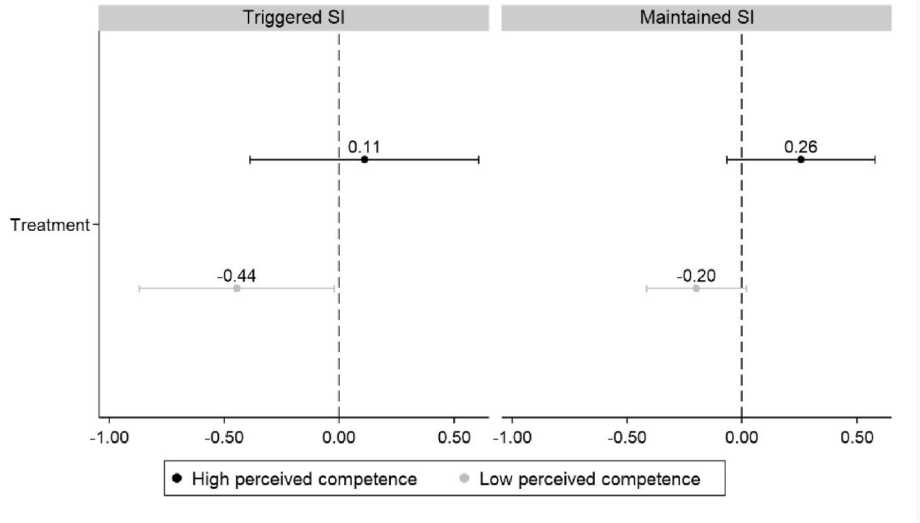


Fig. 5. Subgroup analysis by perceived competence.

the role of example choice in mathematics education during a single learning session. Though this potentially explains the differences in findings, we believe other arguments are worth exploring. In particular, example choice in our study was characterized by shallow to medium depth (some exercises inserted out-of-school interest-related words into templates, while other exercises provided relevant connections to out-of-school interests with an opportunity for contextual grounding), medium grain size (popular out-of-school interest fields were chosen based on a survey with students in the seventh and eighth grades, yet, the contexts were not fully personalized), and a high level of ownership (students were given multiple choice moments). Comparing the depth, grain size, and level of ownership across studies, we find that the depth and grain size were fairly similar. Nevertheless, the level of ownership in our study was much higher compared to the previous studies where the chosen context was held constant during the entire learning session. Similar to the argument made by Clinton and Walkington (2019) for a fine grain size, a high level of ownership may have involved extra information to process, thereby offsetting some of the benefits for learning. In particular, offering more choices was potentially unattractive due to judgments that involved additional information (Schwartz, 2000), leading to a cognitive burden (Patall et al., 2008). In turn, the cognitive burden may have had disruptive effects on students' sense of personal autonomy and motivation (de Brabander & Martens, 2014). Hence, the findings of our study suggest that a greater degree of choice does not lead to an improved learning experience for students. We contribute to the limited body of evidence for the extent to which example choice can be used throughout a learning session (related to financial education in particular) to hold students' attention (Reber et al., 2009).

Several limitations and future directions need to be acknowledged in our study. First, we do not know how the individual interest of students in the topic evolved over time as we do not have a post-treatment measurement. Ideally, and considering the four-phased, sequential model of interest (Hidi & Renninger, 2006), triggered situational interest, maintained situational interest, and individual interest should be measured at different points in time. Although the individual interest of students would have been an interesting outcome variable, including multiple post-treatment tests was beyond the scope of the study, as including additional post-tests would increase the attrition in the experiments. Second, it should thus also be noted that the triggered situational interest of students, which was assessed at the same time as maintained situational interest, was less sensitive than those used in previous studies (Høgheim & Reber, 2015; 2017). Third, we acknowledge the difficulty of the learning material (which partially explained the high non-compliance) as a limitation of our study. Fourth, we do not have data on how students approached the task. In particular, we do not know how much time students spent on the exercises (i.e., there were no timestamps available in the computer-assisted program) and whether students in the treatment condition consistently chose a context that was aligned with their out-of-school interests or switched between interest fields, including those not aligned with their out-of-school interests. These data could provide evidence consistent with cognitive overloading. Fifth, to guarantee adequate sample sizes per condition, our study compared a computer-assisted program including an extended form of choice and personalization with a computer-assisted program without choice or personalization. To isolate the effects of choice and personalization, future field experiments could include an additional treatment condition where students are not provided a choice, i.e., where the context is held constant during the entire program. Sixth, an issue not addressed in the previous studies, nor in ours, is whether example choice has long-run effects. Although our intervention included multiple learning sessions, the effect of the program was measured shortly after the last lecture. Hence, while example choice in our study did not yield short-run effects on the financial knowledge and interest in the financial topic of students, it may have made students think more about the topic, affecting their performance in the longer run (Reber et al., 2009). This suggests an avenue for further research. Finally, computer-based learning environments appear the most effective when the instructional method includes an element of face-to-face instruction (Tamim et al., 2011). Hence, it might be interesting to examine the effectiveness of example choice in a setting where teachers play a more active role in students' development of interest.

To conclude, the present paper has important implications for educational practice. While many educators have been led to believe that there are large benefits of adjusting material to the out-of-school interests of students, the findings of this study show that an extended form of choice and personalization in a computer-based learning environment does not result in improved student outcomes. Accordingly, educators need to act carefully when considering interest-based interventions. The depth, grain size, and ownership of such interventions are important (Clinton & Walkington, 2019). More research is needed to determine the optimal degree of choice and personalization during computer-assisted learning.

Credit author statement

Kaat Iterbeke: Conceptualization, Investigation, Formal analysis, Writing - Original Draft. Wouter Schelfhout: Writing - Review & Editing,

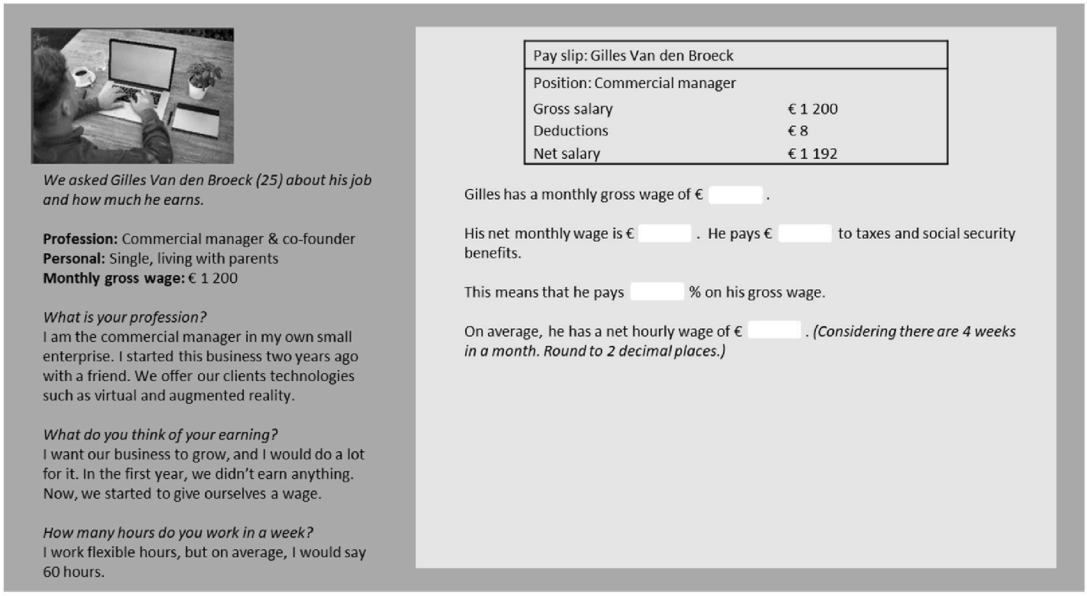
Supervision, Funding acquisition. Kristof De Witte: Conceptualization, Methodology, Writing - Review & Editing, Supervision, Funding acquisition.

Author note

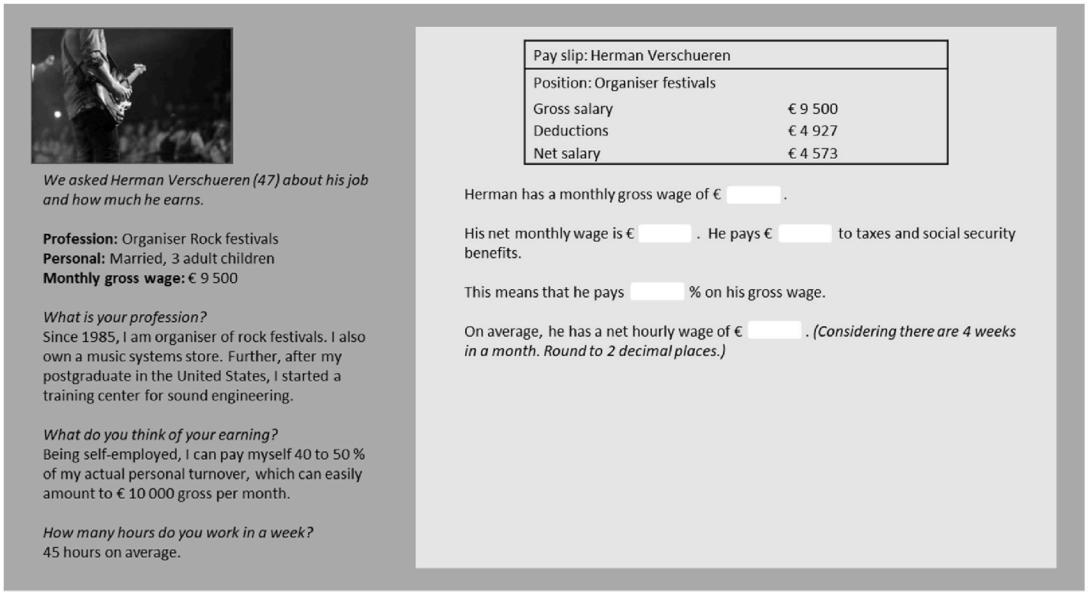
The study was supported by the Flemish Science Organisation [grant S000617N]. The study is registered in the AEA RCT Registry with the unique identifying number AEARCTR-0005480. We thank Lieve Lammens, Johan Mestdagh, Koen Declercq, and Joana Elisa Maldonado for their research assistance. We thank Fenella Carpena and conference participants at LEER-EUFin 2020 and EUFin 2021 for valuable suggestions on an earlier version of this paper.

Appendix

Appendix A. Figures

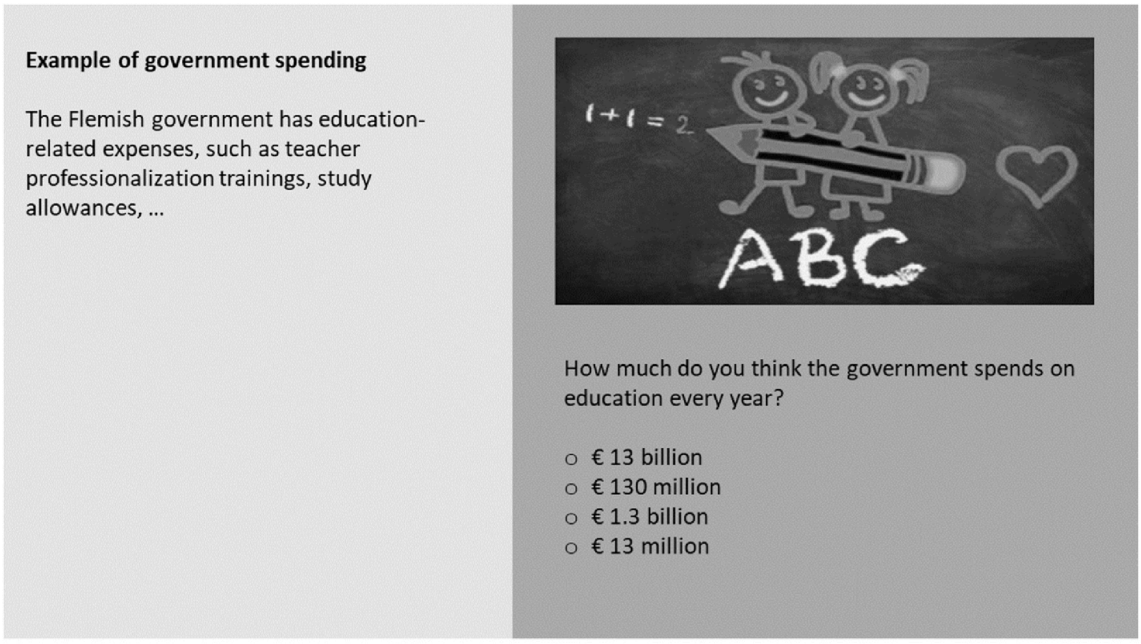


(a) Exercise in the control condition (generic context)



(b) Exercise in the treatment condition (adapted to the context music and culture)

Fig. A.1. Example of Differences in Instruction in the Learning Material.



(a) Exercise in the control condition (generic context)

(b) Exercise in the treatment condition (adapted to the sports-related context)

Fig. A.2. Example of Differences in Instruction in the Learning Material. *Note.* The questions shown above served as an example of government spending. After completing these questions, students in the control and treatment conditions were given the same learning content related to government expenses.

Appendix B. Tables

Table B.1
Descriptive Statistics of Interest Survey

Interest field	Seventh grade (N = 115)	Eighth grade (N = 132)
Sports	30.43%	32.58%
Social media (YouTube, Instagram, …)	23.48%	25.00%
Gaming	13.91%	12.88%

(*continued on next page*)

Table B.1 (*continued*)

Interest field	Seventh grade (N = 115)	Eighth grade (N = 132)
Music and culture	11.30%	13.64%
Movies and books	14.78%	5.30%
Youth movement	6.09%	10.61%

Note: Students were asked to indicate which of the six popular interest fields interested them the most. Note that, when designing the website, we combined the interest fields "Social media" and "Gaming" because of context similarities.

Table B.2
Sample Representativeness

School characteristic	In-sample Schools	Out-of-sample Schools	*p-value*
Panel A. Schools at baseline			
% low educated mothers	25.3	24.5	0.813
% on allowance	30.6	30.1	0.861
% non-native	20.0	17.4	0.465
% neighborhood high retention	28.1	24.7	0.432
Panel B. Schools at follow-up			
% low educated mothers	20.4	24.6	0.322
% on allowance	25.0	30.2	0.183
% non-native	15.9	17.5	0.719
% neighborhood high retention	21.1	24.8	0.517

Note. Mean values and *p*-value of each school characteristic are computed using a *t*-test; There are three missing participating schools in the administrative data.

Table B.3
Non-compliance of Teachers

	Mean	Non-compliance	Treatment	Treatment x Non-compliance
Private education	0.906	−0.132	−0.136	0.201
	(0.0703)	(0.168)	(0.147)	(0.224)
Fraction of 8th grade students	0.948	0.0519	0.0519	−0.194
	(0.0556)	(0.0556)	(0.0556)	(0.121)
Track (Academic)	0.948	−0.00361	−0.0217	−0.0651
	(0.0556)	(0.0779)	(0.0931)	(0.148)
Average baseline test score (15)	6.514	−0.269	−0.101	0.0601
	(0.363)	(0.526)	(0.633)	(0.851)
Teaching blocks (4)	2.595	0.165	0.392	−0.152
	(0.307)	(0.334)	(0.392)	(0.422)

Note. Clustered standard errors at school level in parentheses; ***p < 0.01, **p < 0.05, *p < 0.1; Each row represents a regression of the row variable on the column variables; Non-compliance equals one if the teacher did not assign the post-test in class, zero otherwise; Interaction term indicates whether non-compliance is random across conditions; Value in parentheses after a characteristic refers to maximum value; Values for track refer to pre-vocational education (0) and academic education (1); Teaching blocks refer to the planning of lectures (1 = one block of lecture hours, 2 = two block of lecture hours, 3 = four blocks of lecture hours, 4 = n/a).

Table B.4
Importance of Attrition

	Mean	Attrition	Treatment	Treatment x Attrition
Private education	0.908	−0.125	−0.206	0.289
	(0.0697)	(0.155)	(0.173)	(0.212)
Fraction of 8th grade students	0.947	0.0504	0.0531	−0.145
	(0.0568)	(0.0541)	(0.0568)	(0.0900)
Track (Academic)	0.947	−0.00113	−0.0408	−0.00914
	(0.0568)	(0.0736)	(0.109)	(0.129)
Age	13.03	0.126*	0.0603	−0.0749
	(0.0423)	(0.0687)	(0.0722)	(0.109)
Gender (female)	0.458	0.0487	0.0144	−0.0103
	(0.0523)	(0.0529)	(0.0744)	(0.0889)
Language (Dutch)	0.824	−0.0598	0.0627	0.0181
	(0.0596)	(0.0664)	(0.0811)	(0.0879)
SES (4)	3.140	−0.283**	0.00920	0.0594
	(0.131)	(0.104)	(0.160)	(0.150)
Individual interest (5)	3.331	0.0179	−0.143	−0.0464
	(0.0460)	(0.0759)	(0.129)	(0.148)
Perceived competence in topic (5)	2.701	0.0983**	0.00762	−0.0904
	(0.0631)	(0.0441)	(0.0883)	(0.0939)
Baseline test score (15)	6.472	−0.166	0.0846	−0.224
	(0.389)	(0.517)	(0.653)	(0.734)
Government and taxes (11)	4.701	−0.0697	−0.0377	−0.165
	(0.315)	(0.363)	(0.506)	(0.537)
Financial literacy (4)	1.771	−0.0962	0.122	−0.0586
	(0.0867)	(0.177)	(0.169)	(0.248)

Note. Clustered standard errors at school level in parentheses; ***p < 0.01, **p < 0.05, *p < 0.1; Each row represents a regression of the row variable on the column variables; Attrition equals one if the students did not complete the post-test, zero if the student completed both tests; Interaction term indicates whether attrition is random across conditions; Value in parentheses after a characteristic refers to maximum value; Values for track refer to pre-vocational education (0) and academic education (1); SES is approximated by the number of times a student travels abroad during the year (1 = not, 2 = 1 time, 3 = 2 times, 4 = more than 2 times).

Table B.5
Difference-In-Differences Estimation

Dependent variable	Overall score	Score G&T
Treatment	0.0290	−0.0166
	(0.226)	(0.225)
Time	0.500**	0.515**
	(0.203)	(0.192)
Time * Treatment	−0.162	−0.142
	(0.236)	(0.236)
Observations	1334	1334
R^2	0.043	0.048

Note. Abbreviation G & T refers to the questions related to the government and taxes; Clustered standard errors at school level in parentheses; ***p < 0.01, **p < 0.05, *p < 0.1.

Table B.6
Intent-to-Treat Analysis with Control Variables

Dependent variable	Overall score	Score G&T	Triggered SI	Maintained SI	Perceived value
Treatment	−0.113	−0.128	−0.318	−0.0766	−0.0123
	(0.249)	(0.257)	(0.194)	(0.0999)	(0.114)
Baseline test score	0.459***	0.391***	0.103**	0.121***	0.0705
	(0.0615)	(0.0516)	(0.0395)	(0.0326)	(0.0518)
Individual interest	0.00463	0.0296	0.157***	0.416***	0.358***
	(0.0351)	(0.0442)	(0.0490)	(0.0378)	(0.0525)
Perceived competence	−0.0760*	−0.0556	0.0600	0.0708*	0.0128
	(0.0365)	(0.0351)	(0.0699)	(0.0358)	(0.0429)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	667	667	645	646	646
R^2	0.370	0.322	0.112	0.243	0.180

Note. Abbreviations G&T and SI refer to the questions related to the government and taxes and situational interest, respectively; Clustered standard errors at school level in parentheses; ***p < 0.01, **p < 0.05, *p < 0.1; Baseline test score in column 3–10 refers to the G&T score; Controls: gender, language spoken at home, SES, track, grade, type of education (private/public).

Table B.7
Lee Bounds Analysis

Dependent variable	Overall score		Score G&T		Triggered SI		Maintained SI		Perceived value	
	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Treatment	−0.356	0.0465	−0.358	0.0609	−0.620***	−0.0766	−0.385***	0.131	−0.387***	0.166
	(0.231)	(0.202)	(0.235)	(0.210)	(0.160)	(0.178)	(0.0809)	(0.101)	(0.102)	(0.110)
Baseline test score	0.567***	0.607***	0.520***	0.536***	0.121***	0.0894**	0.136***	0.116***	0.125**	0.0909*
	(0.0626)	(0.0371)	(0.0663)	(0.0384)	(0.0269)	(0.0331)	(0.0293)	(0.0398)	(0.0437)	(0.0442)
Individual interest	−0.0183	−0.0274	0.0209	0.0148	0.128***	0.148***	0.406***	0.439***	0.321***	0.361***
	(0.0368)	(0.0461)	(0.0489)	(0.0566)	(0.0415)	(0.0486)	(0.0306)	(0.0336)	(0.0522)	(0.0433)
Perceived competence	−0.0808**	−0.0865**	−0.0700**	−0.0760***	0.122*	0.131**	0.0873**	0.126***	−0.00407	0.0134
	(0.0288)	(0.0373)	(0.0304)	(0.0255)	(0.0580)	(0.0597)	(0.0325)	(0.0273)	(0.0382)	(0.0418)
Observations	620	620	620	620	588	588	589	589	589	589
R^2	0.344	0.355	0.303	0.290	0.182	0.071	0.320	0.274	0.214	0.173

Note. Abbreviations G&T and SI refer to the questions related to the government and taxes and situational interest, respectively; Clustered standard errors at school level in parentheses; ***p < 0.01, **p < 0.05, *p < 0.1; We estimate the upper and lower bounds by manually trimming the condition that is least affected by attrition (here, the control condition) to equalise response rate for all outcome measures. To generate the trimming fractions, we use the residuals from a regression of the outcome value on control variables with standard errors clustered at school level.

Table B.8
Coarsened Exact Matching Analysis

Dependent variable	Overall score	Score G&T	Triggered SI	Maintained SI	Perceived value
Treatment	−0.174	−0.158	−0.383	−0.105	−0.0155
	(0.236)	(0.237)	(0.226)	(0.124)	(0.115)
Baseline test score	0.530***	0.456***	0.134*	0.200***	0.0992
	(0.0662)	(0.0516)	(0.0678)	(0.0354)	(0.0744)
Individual interest	−0.0209	0.0315	0.185**	0.428***	0.386***
	(0.0752)	(0.0876)	(0.0832)	(0.0436)	(0.0569)
Perceived competence	−0.135**	−0.126**	0.0557	0.0670	−0.0329
	(0.0471)	(0.0528)	(0.0831)	(0.0392)	(0.0582)
Observations	469	469	451	452	452
R^2	0.259	0.196	0.094	0.225	0.141

Note. Abbreviations G&T and SI refer to the questions related to the government and taxes and situational interest, respectively; Clustered standard errors at school level in parentheses; ***p < 0.01, **p < 0.05, *p < 0.1; Baseline test score in column 2–5 refers to the G&T score; Variables used for matching: type of education (private/public), grade, language spoken at home, gender, track, SES, baseline test score, individual interest, and perceived competence.

Table B.9
Wild Cluster Restricted Bootstrap

Dependent variable	Overall score	Score G&T	Triggered SI	Maintained SI	Perceived value
Treatment	−0.151	−0.145	−0.333	−0.110	−0.101
	(0.247)	(0.252)	(0.198)	(0.106)	(0.124)
p-value wild bootstrap	0.707	0.691	0.243	0.414	0.522
Baseline test score	0.583***	0.519***	0.136***	0.156***	0.136***
	(0.0554)	(0.0501)	(0.0370)	(0.0295)	(0.0454)
p-value wild bootstrap	0.000***	0.000***	0.008***	0.003***	0.021**
Individual interest	−0.0134	0.0217	0.139***	0.404***	0.343***
	(0.0349)	(0.0463)	(0.0477)	(0.0346)	(0.0510)
p-value wild bootstrap	0.756	0.718	0.009***	0.000***	0.000***
Perceived competence	−0.0823**	−0.0657**	0.0797	0.0725**	0.000326
	(0.0326)	(0.0309)	(0.0651)	(0.0331)	(0.0381)
p-value wild bootstrap	0.015**	0.033*	0.281	0.034**	0.994
Observations	667	667	645	646	646
R^2	0.303	0.248	0.087	0.226	0.153

Note. Abbreviations G&T and SI refer to the questions related to the government and taxes and situational interest, respectively; Clustered standard errors at school level in parentheses; Small number of (treated) clusters in data are controlled for using the wild cluster restricted bootstrap method; ***p < 0.01, **p < 0.05, *p < 0.1; Baseline test score in column 2–5 refers to the G&T score.

Table B.10
Heterogeneous Effects by Individual Interest

Dependent variable	Overall score	Score G&T	Triggered SI	Maintained SI	Perceived value
Treatment	−0.151	−0.145	−0.333	−0.110	−0.101
	(0.247)	(0.252)	(0.198)	(0.106)	(0.123)
Treatment * Individual interest	−0.00861	0.00824	0.0236	0.0609	0.0628
	(0.0791)	(0.0964)	(0.0996)	(0.0613)	(0.0863)
Baseline test score	0.582***	0.520***	0.136***	0.158***	0.138***
	(0.0557)	(0.0501)	(0.0364)	(0.0290)	(0.0444)
Individual interest	−0.00909	0.0176	0.128*	0.374***	0.312***
	(0.0398)	(0.0475)	(0.0703)	(0.0423)	(0.0646)
Perceived competence	−0.0824**	−0.0656**	0.0799	0.0728**	0.000702
	(0.0324)	(0.0307)	(0.0651)	(0.0325)	(0.0376)
Observations	667	667	645	646	646
R^2	0.303	0.248	0.087	0.227	0.154

Note. Abbreviations G&T and SI refer to the questions related to the government and taxes and situational interest, respectively; Clustered standard errors at school level in parentheses; ***p < 0.01, **p < 0.05, *p < 0.1; Baseline test score in column 2–5 refers to the G&T score.

Table B.11
Heterogeneous Effects by Perceived Competence

Dependent variable	Overall score	Score G&T	Triggered SI	Maintained SI	Perceived value
Treatment	−0.153	−0.146	−0.324	−0.105	−0.0989
	(0.247)	(0.252)	(0.190)	(0.105)	(0.123)
Treatment * Perceived competence	−0.0708	−0.0354	0.273**	0.155***	0.0769
	(0.0591)	(0.0571)	(0.0997)	(0.0496)	(0.0685)
Baseline test score	0.583***	0.519***	0.136***	0.157***	0.136***
	(0.0553)	(0.0499)	(0.0363)	(0.0300)	(0.0456)
Individual interest	−0.0138	0.0215	0.140***	0.404***	0.343***
	(0.0351)	(0.0464)	(0.0478)	(0.0352)	(0.0511)

(*continued on next page*)

Table B.11 (*continued*)

Dependent variable	Overall score	Score G&T	Triggered SI	Maintained SI	Perceived value
Perceived competence	−0.0515	−0.0503	−0.0375	0.00567	−0.0327
	(0.0395)	(0.0374)	(0.0372)	(0.0200)	(0.0335)
Observations	667	667	645	646	646
R^2	0.304	0.248	0.105	0.232	0.155

Note. Abbreviations G&T and SI refer to the questions related to the government and taxes and situational interest, respectively; Clustered standard errors at school level in parentheses; ***p < 0.01, **p < 0.05, *p < 0.1; Baseline test score in column 2–5 refers to the G&T score.

Table B.12
Heterogeneous Effects by Baseline Test Score

Dependent variable	Overall score	Score G&T	Triggered SI	Maintained SI	Perceived value
Treatment	−0.151	−0.145	−0.334	−0.110	−0.101
	(0.247)	(0.252)	(0.197)	(0.106)	(0.123)
Treatment * Baseline test score	0.0385	0.0107	−0.0403	0.0488	−0.0129
	(0.108)	(0.102)	(0.0613)	(0.0549)	(0.0874)
Baseline test score	0.565***	0.514***	0.154***	0.134***	0.142**
	(0.0796)	(0.0770)	(0.0518)	(0.0337)	(0.0534)
Individual interest	−0.0123	0.0220	0.138***	0.405***	0.342***
	(0.0354)	(0.0464)	(0.0470)	(0.0339)	(0.0499)
Perceived competence	−0.0823**	−0.0657**	0.0797	0.0725**	0.000317
	(0.0327)	(0.0309)	(0.0651)	(0.0328)	(0.0382)
Observations	667	667	645	646	646
R^2	0.304	0.248	0.087	0.227	0.153

Note. Abbreviations G&T and SI refer to the questions related to the government and taxes and situational interest, respectively; Clustered standard errors at school level in parentheses; ***p < 0.01, **p < 0.05, *p < 0.1; Baseline test score in column 2–5 refers to the G&T score.

Appendix C. Post-treatment Test

Item	Reference
Knowledge	
1. What is the government not responsible for?	Learning material
o Ensure a fair income distribution	
o Provide services for all citizens, such as the army	
o Provide services for all citizens, such as the police	
o Provide services for all citizens, such as free public transport	
o I don't know the answer	
2. Which authority is responsible for culture in Belgium?	Learning material
o The federal government	
o The regional government	
o The municipalities	
o I don't know the answer	
3. Who is in charge of justice?	Learning material
o The federal government	
o The regional government	
o The municipalities	
o I don't know the answer	
4. Which of the following examples is an expense of the government?	Learning material
o Corporate tax	
o Excises	
o Registration tax	
o Unemployment benefits	
o I don't know the answer	
5. Which of the following examples is a revenue of the government?	Learning material
o Interest on public debt	
o Child benefit	
o Pension	
o Inheritance tax	
o I don't know the answer	
6. Every month, Hannah's employer pays money to Hannah's bank account. This is her pay slip for July.	OECD (2016)

Employee Pay slip: Hannah	
Position: manager	1 July to 31 July
Gross salary	€ 3 000
Deductions	€ 1 000
Net salary	€ 2 000
Gross salary to date this year	€ 24 000

(*continued on next page*)

(*continued*)

Item						Reference
	How much money did Hannah's employer pay to her bank account on 31 July? o € 3000 o € 1000 o € 2000 o € 24 000 o I don't know the answer					
7.	Eva has a gross wage of € 2000 per month and a net wage of € 1400 per month. How much taxes and social security contribution does she pay on her gross wage? o 14% o 30% o 60% o 600% o I don't know the answer					Learning material
8.	Lisa is a single mom of two children. As a fulltime child-minder, her gross wage is € 1650 per month. What benefit can Lisa not receive from the government in her current situation? o A replacement income o A supplementary income o Child benefit o Reimbursement of healthcare costs o I don't know the answer					Learning material
9.	Sebastian has two children. He works fulltime as a solicitor and earns € 3500 net per month. Which benefit can Sebastian receive from the government in his current situation? o An unemployment benefit o Social housing o Reimbursement of healthcare costs o A living wage o I don't know the answer					Learning material
10.	A good redistribution of income ensures: o More criminality o Less poverty o A worse healthcare system o A worse education system o I don't know the answer					Learning material
11.	Some people want the government to have more responsibilities, whereas others want the government to intervene as little as possible. What is not considered an essential government task according to the supporters of a "small" government? o Safety of citizens o Safety of belongings o Healthcare and education o Justice o I don't know the answer					Learning material
12.	Laura borrows € 2500 from the bank at an interest rate of 3% to buy a new car. Which of the following statements is correct? o Laura will receive 3% interest from the bank o Laura will pay the bank € 2500 and 3% interest per year on the amount due o Laura will pay the bank € 2500 o I don't know the answer.					Lusardi and Mitchell (2011)
13.	You open up a savings account and deposit € 200. The interest on the savings account amounts to 2% per year. How much money will be on your savings account after five years, if you do not withdraw or deposit additional amounts: o Less than € 220 o Exactly € 220 o More than € 220 o I don't know the answer					Lusardi and Mitchell (2011)
14.	You open up a savings account and deposit € 50. The interest on the savings account amounts to 2% per year. The inflation amounts to 1% per year. After one year, you can buy: o More than today o Less than today o As much as today o I don't know the answer					Lusardi and Mitchell (2011)
15.	Investor A has invested a certain amount of money in shares of one company, whereas investor B the same amount, but spread over shares of multiple companies. Who has the largest risk? o Investor A o Investor B o Both investors have the same risk o I don't know the answer					Lusardi and Mitchell (2011)
Triggered situational interest						
16.	liked the material.					Høgheim & Reber, 2015; Linnenbrink-Garcia et al., 2010
	Totally disagree ○	Disagree ○	No opinion ○	Agree ○	Totally agree ○	
17.	The material caught my attention.					Høgheim & Reber, 2015; Linnenbrink-Garcia et al., 2010
	Totally disagree ○	Disagree ○	No opinion ○	Agree ○	Totally agree ○	
18.	I found the material boring. *(reversed scored)*					Høgheim & Reber, 2015; Linnenbrink-Garcia et al., 2010
	Totally disagree ○	Disagree ○	No opinion ○	Agree ○	Totally agree ○	
19.	The material was fun to work with.					Høgheim & Reber, 2015; Linnenbrink-Garcia et al., 2010
	Totally disagree	Disagree	No opinion	Agree	Totally agree	

(*continued on next page*)

(*continued*)

	○	○	○	○	○	
Maintained situational interest						
20.	I found the topic of government and taxes boring. *(reversed scored)*					Høgheim & Reber, 2015; Linnenbrink-Garcia et al., 2010
	Totally disagree	Disagree	No opinion	Agree	Totally agree	
	○	○	○	○	○	
21.	I liked what I learned.					Høgheim & Reber, 2015; Linnenbrink-Garcia et al., 2010
	Totally disagree	Disagree	No opinion	Agree	Totally agree	
	○	○	○	○	○	
22.	I thought the topic of government and taxes was interesting.					Høgheim & Reber, 2015; Linnenbrink-Garcia et al., 2010
	Totally disagree	Disagree	No opinion	Agree	Totally agree	
	○	○	○	○	○	
23.	I did not like what I learned. *(reversed scored)*					Høgheim & Reber, 2015; Linnenbrink-Garcia et al., 2010
	Totally disagree	Disagree	No opinion	Agree	Totally agree	
	○	○	○	○	○	
Perceived value						
24.	What I learned was useful for me.					Høgheim & Reber, 2015; Linnenbrink-Garcia et al., 2010
	Totally disagree	Disagree	No opinion	Agree	Totally agree	
	○	○	○	○	○	
25.	I think this is valuable to learn.					Høgheim & Reber, 2015; Linnenbrink-Garcia et al., 2010
	Totally disagree	Disagree	No opinion	Agree	Totally agree	
	○	○	○	○	○	
26	What I learned is completely useless. *(reversed scored)*					Høgheim & Reber, 2015; Linnenbrink-Garcia et al., 2010
	Totally disagree	Disagree	No opinion	Agree	Totally agree	
	○	○	○	○	○	
27	I could have done this again since it is valuable to me.					Høgheim & Reber, 2015; Linnenbrink-Garcia et al., 2010
	Totally disagree	Disagree	No opinion	Agree	Totally agree	
	○	○	○	○	○	

References

Bernacki, M. L., & Walkington, C. (2018). The role of situational interest in personalized learning. *Journal of Educational Psychology, 110*(6), 864–881.

Blackwell, M., Iacus, S., King, G., & Porro, G. (2009). CEM: Coarsened exact matching in Stata. *STATA Journal, 9*(4), 524–546.

de Brabander, C. J., & Martens, R. L. (2014). Towards a unified theory of task-specific motivation. *Educational Research Review, 11*, 27–44.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics, 90*(3), 414–427.

Clinton, V., & Walkington, C. (2019). Interest-enhancing approaches to mathematics curriculum design: Illustrations and personalization. *The Journal of Educational Research, 112*(4), 495–511.

Cordova, D. I., & Lepper, M. R. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology, 88*(4), 715–730.

Flowerday, T., Schraw, G., & Stevens, J. (2004). The role of choice and interest in reader engagement. *The Journal of Experimental Education, 72*(2), 93–114.

Ghysels, J., Haelermans, C., & Prince, F. (2014). *The economics of information in human capital formation— evidence from two randomized experiments with low stakes tests in secondary education (TIER WP 14/27), Maastricht* (TIER Working Paper Series).

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology, 90*(3), 414–434.

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111–127.

Høgheim, S., & Reber, R. (2015). Supporting interest of middle school students in mathematics through context personalization and example choice. *Contemporary Educational Psychology, 42*, 17–25.

Høgheim, S., & Reber, R. (2017). Eliciting mathematics interests: New directions for context personalization and example choice. *The Journal of Experimental Education, 85*(4), 597–613.

Iacus, S., King, G., & Porro, G. (2008). Causal inference without balance checking: Coarsened exact matching. *Political Analysis, 20*(1), 1–24.

Iterbeke, K., De Witte, K., Declercq, K., & Schelfhout, W. (2020). The effect of ability matching and differentiated instruction in financial literacy education. Evidence from two randomised control trials. *Economics of Education Review, 78*, 101949.

Kaiser, T., & Menkhoff, L. (2020). Financial education in schools: A meta-analysis of experimental studies. *Economics of Education Review, 78*, 101930.

Katz, I., & Assor, A. (2007). When choice motivates and when it does not. *Educational Psychological Review, 19*, 429–442.

Koch, A., Nafziger, J., & Skyt Nielsen, H. (2015). Behavioral economics of education. *Journal of Economic Behavior & Organization, 115*, 3–17.

Lai, F., Zhang, L., Qu, Q., Hu, X., Shi, Y., Boswell, M., & Rozelle, S. (2015). Teaching the language of wider communication, minority students, and overall educational performance: Evidence from a randomized experiment in qinghai province, China. *Economic Development and Cultural Change, 63*(4), 753–776.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies, 76*, 1071–1102.

Lehman, S., Schraw, G., McCrudden, M. T., & Hartley, K. (2007). Processing and recall of seductive details in scientific text. *Contemporary Educational Psychology, 32*, 569–587.

Linnenbrink-Garcia, L., Durik, A. M., Conley, A. M. M., Barron, K. E., Tauer, J. M., Karabenick, S. A., & Harackiewicz, J. M. (2010). Measuring situational interest in academic domains. *Educational and Psychological Measurement, 70*(4), 647–671.

Linnenbrink-garcia, L., Patall, E. A., & Messersmith, E. E. (2013). Antecedents and consequences of situational interest. *British Journal of Educational Psychology, 83*, 591–614.

Lusardi, A., & Mitchell, O. S. (2011). *Financial literacy and planning: Implications for retirement wellbeing*. NBER Working Paper Series No 17078.

Magner, U. I. E., Schwonke, R., Aleven, V., Popescu, O., & Renkl, A. (2014). Triggering situational interest by decorative illustrations both fosters and hinders learning in computer-based learning environments. *Learning and Instruction, 29*, 141–152.

Muralidharan, K., Singh, A., & Ganimian, A. J. (2019). Disrupting education? Experimental evidence on technology-aided instruction in India. *The American Economic Review, 109*(4), 1426–1460.

OECD. (2016). *PISA 2015 financial literacy framework. PISA 2015 assessment and analytical framework: Science, reading, mathematic and financial literacy*. Paris: OECD Publishing.

Patall, E. A. (2013). Constructing motivation through choice, interest, and interestingness. *Journal of Educational Psychology, 105*(2), 522–534.

Patall, E. A., Cooper, H., & Robinson, J. C. (2008). The effects of choice on intrinsic motivation and related outcomes: A meta-analysis of research findings. *Psychological Bulletin, 134*(2), 270–300.

Patall, E. A., Sylvester, B. J., & Han, C. (2014). The role of competence in the effects of choice on motivation. *Journal of Experimental Social Psychology, 50*, 27–44.

Reber, R., Hetland, H., Chen, W., Norman, E., & Kobbeltvedt, T. (2009). Effects of example choice on interest, control, and learning. *The Journal of the Learning Sciences, 18*, 509–548.

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

Schneider, S., Nebel, S., Beege, M., & Rey, G. D. (2018). The autonomy-enhancing effects of choice on cognitive load, motivation and learning with digital media. *Learning and Instruction, 58*, 161–172.

Schwartz, B. (2000). Self-Determination: The tyranny of freedom. *American Psychologist, 55*, 79–88.

Subban, P. (2006). Differentiated instruction: A research basis. *International Education Journal, 7*(7), 935–947.

Sweller, J. Merrienboer, JJGVan, Paas, F. G. W. C.. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251–296.

Tamim, R. M., Bernard, R. M., Borokhovski, E., Abrami, P. C., & Schmid, R. F. (2011). What forty years of research says about the impact of technology on learning : A

second-order meta-analysis and validation study. *Review of Educational Research, 81*(1), 4–28.

Tomlinson, C. A., Brighton, C., Hertberg, H., Callahan, C. M., Moon, T. R., Brimijoin, K., Conover, L., & Reynolds, T. (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.

Walkington, C. A. (2013). Using adaptive learning technologies to personalize instruction to student interests: The impact of relevant contexts on performance and learning outcomes. *Journal of Educational Psychology, 105*(4), 932–945.