Improving students' financial literacy by training teachers using an online professional development module

Improving students' financial literacy by training teachers using an online professional development module[1]

Boukje Compen[ab*], Kristof De Witte[bc], Koen Declercq[d] , Wouter Schelfhout[a]

*Corresponding author. E-mail address: boukje.compen@uantwerpen.be
[a] Antwerp School of Education, Faculty of Social Sciences, University of Antwerp
[b] Leuven Economics of Education Research, Faculty of Economics and Business, KU Leuven
[c] UNU-MERIT, Maastricht University
[d] CEREC, UCLouvain - Saint-Louis Bruxelles

Abstract. We evaluate whether a scalable online teacher professional development (OTPD) module that requires little time investment enhances students' financial literacy. Two randomised controlled trials were performed, with 1827 students, 53 teachers and 47 schools participating. The financial education programme on its own increased students' financial knowledge, but did not improve financial behaviour. Regarding the OTPD effects, we observed that students' knowledge scores did not significantly improve, but that behaviour scores were enhanced relative to students whose teachers did not receive access to the OTPD module. In comparison with students in the control condition, behaviour scores improved with 0.39 SD.

Keywords. Online teacher professional development; Financial literacy education; Randomised controlled trial

1. Introduction

[1] We thank the editor and two anonymous reviewers for their valuable suggestions. Furthermore, we are grateful to Wikifin.be, Els Lagrou, Lieve Lammens, and Johan Mestdagh for assistance with the practicalities of this study. In addition, we thank conference participants of ESPE (Bath), EALE (Uppsala), M-BEES (Maastricht), LEER (Leuven), JURE (Antwerp), ORD (Heerlen), and the MU-CEN group (Maastricht). This work is supported by the Research Foundation Flanders (FWO) through the 'Financial Literacy @ School' programme (grant number S000617N) and is registered in the AEA RCT Registry: AEARCTR-000 3501. Declarations of interest: none. Data will be made available after acceptance.

Financial literacy is increasingly perceived as an essential competence for lifelong learning, as proven by its integration into multiple frameworks for twenty-first century skills (Partnership for 21st Century Learning, 2019; World Economic Forum, 2015). The Organisation for Economic Co-operation and Development (OECD) defines financial literacy as 'the knowledge and understanding of financial concepts and risks, and the skills, motivation and confidence to apply such knowledge and understanding in order to make effective decisions across a range of financial contexts, to improve the financial well-being of individuals and society, and to enable participation in economic life' (OECD, 2017, p. 24). Existing research indicates that financial literacy correlates with financial outcomes in both the short and long terms (Hastings et al., 2013). Financially literate consumers invest more frequently in the stock market and engage more frequently in retirement planning, leading to increased wealth accumulation (Lusardi & Mitchell, 2007; van Rooij et al., 2011, 2012). In contrast, low financial literacy is correlated with disadvantageous financial behaviours and outcomes such as high-cost borrowing and debt accumulation (Lusardi & Tufano, 2015; Pak, 2018; Sevim et al., 2012). Moreover, it is argued that the financial literacy of consumers encourages competition and innovation in financial markets, which enhances market efficiency (Hastings et al., 2013; Nicolini, 2019). Given the increasing importance of financial literacy, it is worrisome that only one-third of adults worldwide are financially literate (Klapper et al., 2015).

Striving to enhance the populations' financial literacy levels, an increasing number of nations around the world have developed a strategy to offer school-based financial education (OECD, 2016). Among other reasons for targeting schools, school-based programmes ensure that all students are reached, independent of their socioeconomic background, the capacities of their parents to transfer financial skills, or self-selection in particular courses (Van Campenhout et al., 2017). A recent meta-analysis on the effectiveness of financial education demonstrated a positive and significant effect of school-based financial education (Kaiser & Menkhoff, 2020). Despite shedding light on the various factors that influence the effect sizes of programmes (e.g. duration), the review did not evaluate the influence of the quality and instructional behaviours of teachers, or the impact of teacher training. However, based on a comparison of two financial education programmes in the US, Urban et al. (2020) suggest that the level of teacher training may explain the observed difference in effect sizes. They even conclude that teacher preparation may be critical to the effectiveness of financial education. Therefore, the aim of the present paper is to provide further insight into the role of teacher professional development (TPD) in the effectiveness of financial literacy education.

Competent teachers are indeed considered one of the major prerequisites for effective financial education (Blue et al., 2014). However, current in-service teachers tend to lack confidence concerning their competences on financial topics (De Beckker et al., 2019a; Way & Holden, 2010). While multiple financial education programmes have integrated TPD efforts to prepare teachers

(e.g., Batty et al., 2015; Bruhn et al., 2016; Walstad et al., 2010), the majority of studies have only evaluated student learning outcomes, neglecting to assess the extent to which TPD had contributed to the learning process (Compen et al., 2018). Some exceptions are the studies by Swinton et al. (2007) and Harter and Harter (2012), which showed a beneficial impact on student performance of teachers engaging in a graduate course in personal finance, and participation in teacher training related to Financial Fitness For Life programmes. However, teachers were not randomly assigned to the treatment conditions in either of these studies. In contrast, in the study of Compen et al. (2020), a RCT was performed to evaluate the impact of an online teacher professional development (OTPD) initiative in the form of an interactive webinar series. It was shown that the initiative successfully enhanced students' financial literacy levels. A limitation of the webinar series, however, was its lack of scalability. Only a maximum of twelve teachers could participate simultaneously, since the webinars were hosted by a moderator and since active and collaborative learning was encouraged.

The present study contributes to the existing literature on teacher professional development and financial literacy by conducting a RCT that evaluates whether a highly scalable TPD initiative in the form of OTPD module enhances the effectiveness of a financial education programme on students' financial literacy. This feature of scalability is particularly desirable in contexts of curriculum reform as all teachers can be perceived as novices regarding the expected changes and therefore need adequate training (Starkey et al., 2009; Vivian et al., 2014). Furthermore, as TPD often needs to take place in a relatively short time period in reform contexts, the OTPD module was designed so that it could be covered in approximately three hours. To increase the scalability of the intervention, and in contrast to the webinar series evaluated in Compen et al. (2020), the learning activities in the module were asynchronous and self-directed by the teachers (i.e. teachers could take on the activities at their convenience). Furthermore, as the specific content might matter, it should be noted that earlier work by Compen et al. (2020) evaluated the impact of OTPD on the effectiveness of a financial education programme on saving and investment, while the programme that forms the basis for the present study focusses on the topic of payment methods.

Recently, as part of a large research project funded by the Flemish Science Organisation, a number of randomized controlled trials assessed the impact of introducing financial education in secondary education in the region of Flanders (Belgium). Specifically, Maldonado et al. (2021) and Maldonado and De Witte (2021) evaluated how parental involvement through homework and information provision, respectively, impacts student financial literacy. In Iterbeke et al. (2020a) and Iterbeke et al. (2020b), it was studied to what extent elaborate feedback and differentiated instruction can enhance learning outcomes in the context of financial education. In contrast to the present study, as well as in the study by Compen et al. (2020) that was discussed above, none of the

teachers in these RCTs received any form of training targeted to the provision of financial education. As the topic of payment methods takes a pivotal role in the framework of financial literacy, the present study shares this focus with Maldonado et al. (2021) and Iterbeke et al. (2020a). However, the research question and experimental conditions in the present paper are clearly distinctive from these previous studies. [2]

To address our research objective, we performed two randomised controlled trials among two cohorts of secondary school students in Flanders (Belgium). A total of 1827 students, 53 teachers and 47 schools participated. Schools were randomly allocated to different experimental conditions, and the financial literacy of students was measured before and after the intervention. Students and teachers in the control condition did not have access to the educational material or OTPD but completed the tests at the same time as those in the treatment conditions. In the first treatment condition, students followed a four-hour programme on financial literacy. Comparing the learning outcomes between the control condition and this first treatment condition allowed us to assess the impact of the programme. To gain insight into the potential of OTPD, we added a second treatment condition in which the students followed the programme and their teachers received access to an OTPD module that provided them with more information on the programme's topic.

Our results indicate that the financial education programme increases students' overall financial literacy scores by 0.20 standard deviations (SD) relative to the control condition. This effect was mainly driven by improved financial knowledge scores. Furthermore, we show that the OTPD module is effective in enhancing the effect of the programme on students' financial literacy levels. Students of teachers who received access to the professional development initiative outperform students in the control condition by 0.39 SD when focusing on the financial behaviour scores, resulting in a significant difference in the effects of the two treatment conditions. Although we did not find a significant effect of OTPD on students' financial knowledge scores, students' overall financial literacy levels improved due to the OTPD. These effects should be interpreted as intention-to-treat (ITT) effects as (1) participation in the OTPD module was on a voluntary basis for teachers assigned to this treatment condition, and (2) because teachers in both treatment conditions could deviate from the instructions that were provided. While we cannot control for this second source of potential endogeneity, we control for non-random selection into the OTPD module by making use of information on teachers' actual participation (i.e. consulting the module at least once) and the intensity of treatment (i.e. the number of clicks in the module). We follow an instrumental variables

[2] Also the studies by Bargagli-Stoffi et al. (2021), Compen et al. (2018) and Compen and Schelfhout (2021) stem from the same research project on financial education in Flanders. However, in these studies using machine learning and qualitative methods, other research designs than RCTs were performed to obtain the results.

(IV) approach to identify an average treatment effect on the treated for students whose teachers actively participated in the OTPD module relative to students whose teachers did not participate in the OTPD module. The causal impact of participation and intensity of treatment can be identified by instrumenting both measures of teacher participation with assignment to the OTPD condition. The results of both regressions confirm the ITT estimates in showing that participation in the OTPD module enhances students' learning outcomes. Underlying mechanisms for the obtained effects may be the observed differences in teachers' self-efficacy regarding the provision of financial literacy education, as well as the differences in the duration of the introduction of the programme that were identified based on classroom observations.

The remainder of this article is organised as follows. Section 2 provides a description of the experimental design. Section 3 describes the final sample and its baseline characteristics, while the results of the empirical analyses are presented in Section 4. Section 5 provides insights into the underlying mechanisms of the findings, and Section 6 ends with some concluding remarks.

2. Experimental design

2.1. Context of study

This study was performed in Flanders, the Dutch-speaking region of Belgium. A secondary education trajectory is followed by students typically aged between 12 and 18 years, and consists of three cycles, each composed of two learning years. In the first cycle, students follow either a pre-vocational or a more theoretical track, but choose electives to prepare for the specific track they will eventually follow. Each school is part of one of three educational networks. These networks rely on a similar type of funding and are organised by either the Flemish Community, by city districts/municipalities, or by private organisations. Schools belonging to the first two educational networks are referred to as public schools. Most private schools are Catholic schools, which are highly regarded and therefore tend to attract students of relatively high socio-economic status. The Flemish curriculum and the corresponding learning standards are determined by the government and do not differ between the three types of schools. The school networks and the individual schools have a certain autonomy in how these standards are addressed. Furthermore, there are no centralised tests that students should pass to continue to the next year.

Compared to other OECD regions, 15-year-olds in Flanders score well on PISA assessments of financial literacy. Specifically, the average score of Flemish students is 541, compared to the OECD average of 489. This implies that solely the region of Shanghai-China outperforms Flanders (OECD, 2017). However, 12% of the Flemish students do not reach the baseline score. Furthermore, there

has been an increase of three percentage points in the share of students not reaching the baseline score compared to the PISA test administered in 2012.

Following other countries and regions in Europe, educational reform mandating education on financial topics in all Flemish schools was implemented in September 2019. Since our intervention took place at the beginning of 2018, this ensured that other financial education initiatives did not interfere with our experiment.

2.2. Intervention

The intervention consisted of a financial education programme for students and an OTPD module for teachers. The material focused on payment methods and was delivered in the form of an educational game that was played using a digital medium, in which pairs of students solved problems using the information provided in a booklet.

It is increasingly recognised that financial literacy has two dimensions: knowledge (i.e. understanding) and application (i.e. translating this knowledge into appropriate financial behaviour) (Huston, 2010). Following the reasoning of Antonietti et al. (2016), we attempted to develop a programme that was not limited to increasing students' financial knowledge, but would also help them change their behavioural intentions in their daily lives. Based on Lührmann et al. (2015), the content was related to students' everyday financial decision-making. Since incentivising students may increase their efforts (e.g., Angrist & Lavy, 2009), a small prize was provided for the pair of students finishing the game first in their class.

The programme could be integrated in any subject and was designed to cover four lessons of 50 minutes. Teachers decided whether they would teach the lessons separately, in two blocks, or in one block. The material was standardised and designed to limit preparation time for teachers to enhance the scalability of the programme. The educational material was tested in two pilot schools.

As there is increasing awareness of the importance of differentiated instruction, which refers to the tailoring of instructions to meet the varying learning needs of different students (Iterbeke et al., 2020a), this concept was integrated directly into the material. Teachers were instructed to form pairs of students based on the grades they had most recently obtained in mathematics. The pairs of students were divided into three categories based on mathematics scores, with those who had the lowest scores within the class receiving a version of the material that contained hints to solve the problems. Those in the group with the highest scores received a version without hints, while the other group of students received hints, but to a lesser extent. With this procedure, we aimed to make the programme sufficiently challenging for students with relatively strong mathematical abilities, while simultaneously ensuring that the weaker students could comprehend all of the content covered.

The teachers whose schools were randomly assigned to the treatment condition with OTPD (see Section 2.4) received access to an OTPD module. The aim of this module was to enhance teachers' knowledge of payment methods (content knowledge) and of how to apply differentiated instruction in the context of financial education (pedagogical content knowledge). The module provided information in the form of videos, quizzes and discussion forums, as well as other forms. Teachers were not obliged to engage in the module, and they could decide for themselves which sections to view, or which links to additional or external information to open. This allowed us to meet the potentially varying needs of the participating teachers (Chen et al., 2009). To increase scalability and simultaneously reduce non-compliance as much as possible, the environment was developed so that teachers required a maximum of three hours to cover the core aspects.

2.3. Research design

To examine the effects of the OTPD module on the financial literacy levels of students, we conducted two randomised controlled experiments with similar setups. The first experiment ran from January to March 2018 and was intended for students in the second year of secondary school, who are typically aged 13 or 14 (8th grade). The second experiment ran from February to April 2018, targeting students in the third year of secondary school, who are typically aged 14 or 15 (9th grade). The second experiment was conducted to test the validity of the results obtained in the first experiment.

Using an open call by a government-related financial education agency, all schools were eligible to engage in the experiment. The participating students completed a pretest that measured baseline financial literacy. Schools assigned to the treatment conditions received the educational material approximately three weeks after the completion of the pretest.[3] A posttest was administered immediately after students finished the programme.[4] Teachers were instructed to let students complete the tests during school hours to prevent that they would seek help from parents or the internet. Schools were directed to complete the tests and financial education programme within a specified period. As part of Experiment 1, schools received the link to the pretest in mid-January, and taught the programme in February and/or March. In Experiment 2, the schools received the link in the last week of February, and the programme was taught between the second half of April and end of May.

The student tests contained nine questions that were representative of the content covered in the programme. The questions were framed to reflect teenagers' everyday financial decision-

[3] The randomisation took place after registration, but before administration of the pretest. Allocation to one of the conditions was therefore independent of pretest completion.

[4] Six weeks after completion of the posttests, schools received a follow-up test. This test was designed to examine the longer-term effects of the programme. However, the high attrition rate from posttest to follow-up test obstructed the empirical analyses.

making environment. Six questions measured knowledge on the programme's topics, such as recognising different payment methods and calculating discounts. Three questions captured behavioural aspects related to payment methods. In particular, financial behaviour was measured through students' responses to the hypothetical scenario of receiving a suspicious email from their bank, and by letting them indicate which activities they considered to amount to careful handling of their bankcard. Thus, these items measured students' ability to make proper judgements and assessed how they would apply their practical knowledge in particular situations. This conceptualisation differs from previous studies, where students usually self-report on current money management practices such as saving (Kalmi, 2018). Considering that parents – who play an important role in the financial socialisation (Agnew & Cameron-Agnew, 2015; Maldonado et al., 2021) – may still largely take responsibility for the money management of their children at this age, measuring actual current behaviour may not be most informative, especially considering that students may only make a limited number of independent or other financial decisions at this stage of their lives (Kaiser & Menkhoff, 2020). Therefore, we phrased items that allowed us to measure the students' intentions to behave in a particular manner. An overview of the test items constituting the three different scores is presented in Table A1 in Appendix A. Table A2 discusses indicators of the quality of the test instruments.

To be able to control for individual-level predictors of financial literacy (De Beckker et al., 2019b; Riitsalu & Põder, 2016), the pretest also examined several characteristics that have previously been shown to impact financial literacy levels, that is, gender, age, socioeconomic status, the language spoken at home, motivation to learn about financial topics and self-efficacy regarding financial knowledge (e.g., Lusardi et al., 2010; OECD, 2017). The posttest included questions similar to the pretest (i.e. measuring the same concepts but using different wording and different numbers), thus providing a valid measure that could reveal any changes in financial literacy. To prevent the occurrence of learning effects from the pretest, the correct answers were not shared with the students. Pretests and posttests were matched based on anonymous but unique student IDs. It was emphasised to the students that their test scores would not influence their study progress, but merely served research purposes.

The teachers also completed online tests. In addition to examining a variety of background characteristics (e.g. teaching discipline, perceived importance of financial education, and attitude towards online professional learning), the teachers' financial literacy levels were also assessed. Similar to the tests for students, the pretest and posttest for teachers consisted of nine questions: six assessing financial knowledge and three assessing financial behaviour. To prevent ceiling effects, the majority of the knowledge questions were of a higher order than those included in the student tests. An overview of the test items is provided in Table A1 in Appendix A. The posttest items were similar

to those in the pretest. While the pretests for teachers were completed in the same time period as the student tests, the link to the posttests for teachers was sent six weeks after the completion of the programme.

Implementation of the experimental protocols was monitored on several dimensions. First, the OTPD module allowed us to retrieve log files on the activities of individual teachers in the module. These files were used to examine the intensity of treatment, that is, the extent of teacher participation in the module. This was done because teacher engagement is an endogenous factor which depends on aspects such as teacher motivation. A teacher's level of engagement was derived from the number of logins and the number of clicks in the module. Second, we examined whether, despite the standardised programme, teachers differed in their use of the material. To this end, a random sample of 20% of schools in the treatment conditions were visited for observations. It appeared that the duration of teachers' introductions varied. Thus, despite the educational material being offered in a standardised digital format, teachers with access to the OTPD module had the opportunity to apply their acquired knowledge.

2.4. Experimental conditions

Registered schools were randomly allocated to the control condition or to one of two treatment conditions. Randomisation at the school level prevented spill-over effects between conditions as the assigned treatment was similar for all students and teachers within the same school. Schools in the control condition completed the tests at a similar time to those in the treatment conditions, but they could only implement the financial education programme after the administration of the posttests.

In the first treatment condition (denoted by No OTPD), schools implemented the programme, but were not given access to the OTPD module. Comparing the student learning outcomes for the No OTPD condition with those in the control condition allowed us to evaluate the effectiveness of the financial education programme. Teachers in this condition merely received an online manual for the programme, with the educational objectives of the programme stated in terms of financial knowledge and behaviour, followed by instructions on implementing the programme. After a description of the set-up of the material, teachers were given the solutions to all of the questions. Furthermore, teachers received explicit content discussing all topics covered by the educational material.

In the second treatment condition (denoted by OTPD), teachers received access to the OTPD module in addition to receiving the same educational material and teacher manual as the No OTPD condition. Teachers were encouraged but not obliged to engage in the module. This implies that the extent to which teachers voluntarily engaged, is likely to influence the initiative's effectiveness (Matzat, 2013). By comparing the learning outcomes of students in the No OTPD condition and the

OTPD condition, we were able to evaluate whether teacher participation in the module enhances the effect of the programme. Table 1 provides a schematic overview of the research design and timing.

< Table 1 >

3. Sample selection

3.1. Descriptive statistics

Table 2 displays the descriptive statistics of the sample at the level of the student (Panel A), teacher (Panel B) and school (Panel C). Combining the two experiments, data were collected for 1827 students and 53 teachers from 47 schools. We conducted t-tests to examine differences in characteristics between the control condition and the two treatment conditions.[5]

The first row of Panel A in Table 2 suggests that the distribution of the sample in terms of gender is approximately equal in each condition. On average, the participating students are 14 years old. As a proxy for socioeconomic status (SES), students reported the number of holidays abroad in the last year.[6] Students in our sample travelled abroad on average two times in the last year. We note that 35.8% of the students in the No OTPD condition speak a language other than Dutch at home. This implies that the share of non-native students is significantly larger in comparison with the control and OTPD conditions. To measure students' motivation to learn about financial topics, they were asked to indicate the extent to which they agreed with the statement: '*I find it important to have knowledge on financial matters, such as saving, costs, and payment methods*'. With an average score of four out of five, it seems that the average student considered it important to learn about financial topics. Students in the No OTPD condition scored significantly lower on this item, although the absolute differences between conditions are relatively small. The self-efficacy of students was captured by the statement: '*I find that I have sufficient knowledge on financial matters, such as saving, costs, and payment methods'.* The average score of approximately 3.5 out of 5 implies that students were not convinced of their current knowledge on the topic at baseline. Overall, from these data we can conclude that apart from the language spoken at home, students' background characteristics are similar in the three conditions.

With respect to students' baseline financial literacy, reflected in unstandardised measures of financial knowledge, financial behaviour and total scores, a comparison of the pretest scores for the three conditions showed that all three measures scored higher in the OTPD condition than in the

[5] The descriptive statistics for Experiments 1 and 2 separately are presented in Table B1 in Appendix B.

[6] Students were also asked to indicate the educational level of their mother. However, as 700 students in the sample indicated they did not know, we decided to follow Maldonado et al. (2021), and use the number of holidays abroad as a proxy for SES. Using data from the Household Budget Survey (HBS), a representative survey by the Belgian Statistical Office (Statbel), Maldonado et al. (2021), found a Pearson correlation coefficient of 0.310 between total net household income and expenditure for travel abroad.

control and No OTPD conditions. However, these differences were not significant. The final three rows of Panel A show students' unstandardised scores on the posttest. These results show that average knowledge, behaviour and total scores in the No OTPD condition are higher than in the control condition, but the differences in scores between these two conditions are not statistically significant. Furthermore, they indicate that all three scores for the OTPD condition are significantly higher than in the control condition. These results may be a first indication of the effectiveness of OTPD.

Panel B of Table 2 presents the financial literacy test scores of teachers before and after the intervention, and shows that scores substantially improve in both treatment groups, and only marginally improve in the control condition.[7] In addition, this panel shows that 77% of teachers assigned to the OTPD condition actually engaged in the OTPD module. Panel C shows that the majority of schools in our sample are private schools, and that proportionally fewer schools in the no OTPD condition are academic schools compared to the other two conditions.

In line with Jones et al. (2019) we have performed joint balance tests to determine whether the variables in Panels A, B and C (i.e. the student, teacher and school levels), would jointly predict enrolment into the treatment conditions. The p-values resulting from these tests indicate for each panel an imbalance between the conditions. Therefore, the background characteristics were consistently controlled for in all subsequent analyses. Furthermore, multiple robustness tests that take the imbalance between conditions into account were performed (see Section 4.2.3.).

< Table 2 >

While the t-test results in Table 2 allow us to compare the average pretest and posttest scores between conditions, they do not demonstrate whether the conditions differ significantly in the average *change* from pretest to posttest. The plots in Figure 1 reflect the average difference between conditions in the absolute *change* from students' pretest to posttest scores, with a 95% confidence interval. Focusing on the knowledge score, it is apparent that only the difference between the control condition and the OTPD condition is significant. Thus, the change from pretest to posttest is significantly higher in the OTPD condition than in the control condition. The confidence

[7] We conducted a factor analysis on all background characteristics of teachers derived from the pretest (gender, age, years of teaching experience, having a certificate for teaching economics, having experience with financial literacy education, believing in the importance of financial literacy education at school, perceiving oneself as knowledgeable about financial topics, perceiving oneself as competent to teach financial topics, previous participation in TPD related to financial education and interest in participating in TPD related to financial literacy education). The results show that the pretest score, serving as a proxy for teacher quality, is the only factor with a uniqueness value higher than 0.8. Therefore, only this characteristic is included in the empirical analyses.

intervals suggest that the increase in knowledge is not significantly larger in the No OTPD condition than in the control condition. Nevertheless, it should be noted that this effect is significant at the 10% level in the F-tests on which these plots were based. The graph for the behaviour score indicates that all three comparisons lead to significant differences. Thus, we find that the no OTPD condition outperformed the control condition, and the OTPD condition, in turn, outperformed the no OTPD condition. A similar pattern is observed for the total score.

However, note that this comparison of means only presents a first image of the impact of the educational material and the OTPD module. The results should be interpreted with caution, given that student characteristics are not perfectly balanced at baseline, as was shown in the previous section. These differences between conditions are controlled for in the empirical analysis in Section 4.

Finally, the external validity of the sample is addressed in Table B2 Appendix B. The table demonstrates that the sample is representative for Grade 8 (Experiment 1), but that the share of Grade 9 students (Experiment 2) attending public schools is below the Flemish average.

< Figure 1 >

3.2. Attrition

There are two main causes for the imbalance in some student characteristics at baseline, as observed in Section 3.1. First, the randomisation was implemented at the school level. This higher-level randomisation does not automatically result in a balance in characteristics at the lower level. Second, schools dropped out both before and after randomisation. Figure E1 in Appendix E provides a detailed visualisation of the attrition at each stage of the two experiments. Initially, 44 schools replied positively to the open call in Experiment 1, and 32 schools registered to participate in Experiment 2. However, due to attrition, the final sample consisted of 35 schools in Experiment 1 and 12 schools in Experiment 2. Moreover, as we only included students who completed both the pretest and posttest in the analyses, this further reduced the sample size. To test for differential attrition (i.e. whether the level of attrition differed between the conditions), we estimated the following regression, as in Fryer Jr. (2016):

$$(1)\quad Attrition_{i,k} = \tau_0 + \sum_{k=1}^{3} \tau_k Treatment_{i,k} + \eta X_i + \varepsilon_i$$

where $Attrition_{i,k}$ is a dummy variable with a value of 1 when student *i* completed the pretest but not the posttest. The dummy variable $Treatment_{i,k}$ indicates to which of the k conditions the school

of student *i* was allocated, and τ_k estimates the effect of the treatment on attrition. A vector of control variables at student, teacher and school levels, as presented in Table 2, is captured by X_i. Standard errors, indicated by ε_i, are clustered at the school level.

Table 3 shows the results of the attrition analysis for the analysis on the combined sample (i.e. accumulating the samples of Experiments 1 and 2). When controlling for student' characteristics, attrition in both treatment groups is higher (but only significant at the 10% level) than in the control group. This differential attrition might be due to schools in the control condition having an incentive, as they only received the educational material if they completed both tests. In contrast, there was no incentive to complete the posttest in the treatment conditions. The results of the attrition analyses for Experiments 1 and 2 separately are presented in Table B3 in Appendix B.

< Table 3 >

4. Empirical analysis

4.1. Estimation

The statistical analysis was performed in multiple steps. First, we evaluated the effectiveness of the financial education programme and the OTPD module by comparing student posttest scores between the control condition and each of the two treatment conditions. The results should be interpreted as intention-to-treat (ITT) estimates, as some teachers may not have complied with the allocated treatment. We used the following ordinary least squares (OLS) regression to obtain the estimates:

$$(2)\quad Y_{i,k}^{1} = \beta_0 + \sum_{k=1}^{2} \beta_k Treatment_{i,k} + \gamma Y_i^0 + \delta X_i + \varepsilon_i,$$

with $Y_{i,k}^{1}$ reflecting the outcome of student *i* in experimental condition *k* on the posttest (knowledge score, behaviour score, and total score). β_k estimates the effect of being assigned to experimental condition *k*. We controlled for baseline financial literacy scores Y_i^0 as measured in the pretest, and included a vector of control variables X_i at the student, teacher, and school levels. Standard errors ε_i were clustered at school level.

Since teachers in the OTPD condition were not obliged to engage in the online module, not all teachers did so (as illustrated in Table 2). This implies that teacher participation is an endogenous factor which depends on aspects such as teacher motivation. To account for this non-random selection in the treatment, we instrument the intensity of participation in the OTPD module of student *i*'s teacher by exogenous allocation to the OTPD condition in the first stage of the following two-stage least squares (2SLS) analysis:

(3a) $Engagement_i = \beta_0 + \beta_1 OTPD + \beta_2 Y_i^0 + \theta X_i + \varepsilon_i$

(3b) $Y_i^1 = \gamma_0 + \gamma_1 Treatment_i + \gamma_2 \widehat{Engagement_i} + \gamma_3 Y_i^0 + \lambda X_i + \varepsilon_i$

where $Engagement_i$ is used as a proxy for teacher engagement in the online professional development module. In particular, we assessed teacher's actual participation in the module by monitoring their activity using log files. These log files provided information on the number of logins and the number of clicks in the module. The 'Treatment' variable in the second stage of the 2SLS regression represents the average ITT effect of the financial education programme on student achievement for the No OTPD and OTPD conditions combined. We consider this an ITT effect since there could be other sources of endogeneity regarding the implementation of the programme for which this IV approach does not control. By using assignment to the OTPD condition as an instrument for actual participation in the OTPD module, the IV estimator γ_2 can be interpreted as the treatment effect on the treated (TOT) for the students whose teachers actually participated in the OTPD module. Concretely, it measures the extent to which test scores of students improve if their teachers actively participated in the OTPD module relative to the scores of students whose teachers did not actively participate.

4.2. Empirical results

4.2.1. Intention-to-treat (ITT)

The estimates from the ITT regressions, and their 95% confidence intervals, are graphically presented in Figure 2. We estimated the impact of the financial education programme on the standardised posttest scores of students, presenting the estimates for the knowledge, behaviour and total scores separately. This figure shows the estimates for the combined sample, and for the specification in which we controlled for background characteristics at the student, teacher and school levels. The upper line (i.e. No OTPD – Control) presents the estimate that indicates the effect of receiving financial education. The lower line (i.e. OTPD – Control) indicates the combined effect of financial education and assignment to the OTPD condition.

The graphs indicate that the knowledge scores are, on average, 0.17 SD higher for students in the No OTPD condition than for students in the control condition. The confidence interval in Figure 2 includes zero, which implies that this difference is not significant at the 5% level. However, the more detailed ITT estimates in Table 4 indicate that this effect is significant at the 10% level. The graph further shows that allocation to the OTPD condition improves students' knowledge scores relative to scores of students in the control condition. However, we observe in Table 4 that the difference in

coefficients between the No OTPD condition and OTPD condition is not significant. In other words, there is no additional impact of OTPD on the knowledge score of students.

Focusing on the behaviour scores of students, we observe in both Figure 2 and Table 4 that the educational material itself does not significantly improve these scores. However, assignment to the OTPD condition results in an average increase in behaviour scores of 0.39 SD compared to the control condition. Results of the F-test comparing the estimates of both treatment conditions indicate that students in the OTPD condition significantly outperform those in the No OTPD condition.

The results for the total scores indicate an improvement in scores of 0.20 SD for the No OTPD condition. This effect is statistically significant at the 10% level. Allocation of teachers to the OTPD condition resulted in an estimate of 0.43 SD compared to the control condition. These findings indicate that giving teachers access to the OTPD module benefits students' overall financial literacy levels, and that this improvement is due to increased performance on the behaviour items of the financial literacy test rather than on the knowledge items. The estimates for Experiments 1 and 2 separately, as well as the full regression outputs, are presented in Table B4 in Appendix B.

< Figure 2 & Table 4 >

4.2.2. Intensity of treatment

The log files retrieved from the OTPD module revealed that not all teachers who were assigned to the OTPD condition actually participated in it. This implies that the ITT estimates are not the most informative in indicating the effects of *actual engagement* in the OTPD, and probably underestimate the OTPD effect. Table 5 shows the effect of engagement in the OTPD module on students' posttest scores, while simultaneously controlling for the endogeneity caused by self-selection in the OTPD treatment. Two measures of teacher participation were observed. First, we examined whether a teacher logged in to the module at least once. Second, as a measure of intensity of treatment, we recorded how many times each teacher clicked in the module.

Panel A of Table 5 shows the results of the 2SLS analyses for the dichotomous predictor of Login.[8] The first stage of the 2SLS analysis shows a positive relationship between the instrument (being allocated to the OTPD condition) and the endogenous predictor (Login). When allocated to the OTPD condition, engagement in the module increased by 64 percentage points when control variables were included. However, focusing on the OTPD treatment, teacher participation in the module did not lead to a statistically significant improvement in the knowledge scores of students.

[8] Login = 0 for the teachers in all three conditions who did not log in at all; Login = 1 for teachers in the OTPD condition who logged in at least once.

The behaviour scores of students receiving financial education improved by 0.17 SD. Scores increased by an additional 0.39 SD for students whose teachers engaged in the module. This last effect can be interpreted as an average treatment of the treated effect for the students whose teachers actually participated in the OTPD module. This effect is larger than the corresponding ITT estimate because it now measures the impact of actual participation in the OTPD treatment. The total financial literacy scores indicate that students benefit from the educational material, and that this effect is enhanced when teachers engage in the module. These results confirm the ITT analysis in demonstrating that teacher participation in the OTPD module improves students' financial behaviour scores but does not affect their financial knowledge. The effect of OTPD on the total score is therefore driven by the increased scores on behavioural intentions. Table B5 in Appendix B shows the full regression output.

The results are similar when we use the total number of clicks in the module as a measure of intensity of the treatment, as presented in Panel B of Table 7.[9] The second stage results of the 2SLS analysis show that the intensity of treatment has no effect on students' financial knowledge. However, due to a significant impact on the behaviour scores, an additional effect on students' financial literacy was still observed. Specifically, when control variables are included, 100 additional clicks in the module increased students' financial literacy scores by 0.24 SD.

< Table 5 >

4.2.3. Robustness tests and heterogeneity analysis

A series of tests were run to assess the robustness of the ITT estimates. Table 6 assesses the robustness of the ITT estimates. First, as we randomised at the school level, we repeated the ITT analyses using the average school scores instead of the average student scores as the dependent variables in Panel A. We note that the effect of treatment itself appears more convincing than in the main analyses, whereas the opposite holds for the additional effect of assignment to the OTPD condition. Given the relatively low number of clusters, however, these estimates should be interpreted with caution.

Panel B of Table 6 provides the estimates of a difference-in-differences (DID) model as in Lührmann et al. (2015) and compares the change in test scores between the pretest and posttest of students in the treatment conditions relative to students in the control condition. We solely report the treatment effects, i.e. the interaction effects of the test scores with the post-treatment dummy. The effects of the OTPD module on the behaviour score and total score are, although still positive

[9] Multiplied by 100 in order to obtain meaningful and interpretable coefficients.

and significant, smaller than in the ITT. However, when focusing on the difference between the two treatment conditions, we note that the OTPD condition no longer outperforms the OTPD condition. This implies that the DID model deviates from the ITT analysis.

As a third robustness test, we applied the Coarsened Exact Matching (CEM) algorithm, similar to Iacus et al. (2012), to control for the differences in control variables existing at baseline. In matching approaches, particular observations are pruned from the initial sample based on a selection of control variables. Running the empirical analyses on this more restricted, but also more balanced sample, results in more reliable estimates of the causal effects. The findings of the CEM analysis, in which we compared the effects of the separate treatment conditions with the control condition are presented in Panel C of Table 6. The results indicate that students' financial literacy significantly increases by 0.12 *SD* when following the programme, confirming the effectiveness of the educational material. The OTPD module has a beneficial impact compared to both the control condition and the No OTPD condition, suggesting that assignment to the OTPD condition further improves students' financial literacy.

Fourth, to correct for the differential attrition, we estimated the lower and upper bounds of the ITT effects following the procedure of Lee (2009). These bounds correct for the possibility that findings might be biased when attrition in the sample was non-random (Tauchmann, 2014). The intuition behind the Lee Bounds is that observations are trimmed from either of the conditions, ensuring that the share of observations with observed outcomes is equal. The estimates are presented in Panel D of Table 6. Since in Lee Bounds estimations, control variables such as pretest score are not included in the analysis, we conducted OLS regressions without the inclusion of control variables or the clustering of standard errors. The results indicate that the confidence intervals for the OTPD condition do not contain zero and that the lower bounds are significantly different from zero, indicating that the differential attrition does not lead to a biased result on the impact of the OTPD module. Note that the lower bound for the No OTPD condition is significantly negative. Therefore, the Lee Bounds analysis cannot confirm the finding that the financial literacy education programme itself has a positive effect on the total financial literacy score**.**

We correct for multiple testing using the two-stage procedure discussed in Benjamini et al. (2006) and Anderson (2008). All *p*-values relating to the ITT estimates included in Tables 4 and B4 were transformed into *q*-values. The results in Panel E of Table 6 indicate that the initial ITT estimates are robust when the estimates for the effect of the OTDP module are concerned and could, therefore, not be considered false positives resulting from multiple testing.

Finally, since this study is conducted on three hierarchical levels (i.e. students, classes, and schools), we test whether our results are robust to two-way clustering. Panel F of Table 6 shows that

the standard errors are only limitedly influenced by the type of clustering, leaving the conclusions unaffected.

< Table 6 >

In addition to these robustness tests, a heterogeneity analysis was run to observe whether the OTPD initiative was particularly beneficial to students in specific subgroups. The results in Table 7 show that the effectiveness of the OTPD module does not depend on students' pretest scores or gender. Furthermore, we demonstrate that particularly for high SES students, there is an additional impact of the OTPD module. Interestingly, we observe that the programme itself was shown to be more effective for low SES students than for high SES students. Finally, concerning whether or not the student is a native Dutch speaker, the results indicate that the OTPD module only seems to benefit native Dutch speakers. Similar to the SES results, following the financial education itself seemed to mainly improve the scores of non-native students.

< Table 7 >

5. Mechanisms at the teacher level

We evaluate four mechanisms that could have caused the observed positive effect of the OTPD module. The existing literature suggests that effective TPD results in better teacher quality and that this can ultimately benefit student scores (e.g., Desimone, 2009; Merchie et al., 2016). Therefore, as a first potential mechanism, we examined whether the teachers who engaged in OTPD had higher average posttest scores on the teacher test than those who did not. The results are presented in Panel A of Table 8, from which it is apparent that despite a relatively high average pretest score in the control condition, the teachers in the OTPD condition have the highest absolute average posttest score. However, as the increase in test scores between the pretest and posttest is similar in the two treatment groups, we cannot confirm that the positive effects of the OTPD initiative result from an increase in the teachers' financial literacy.

The second potential underlying mechanism is an increase in teacher efficacy, defined as 'a teacher's judgment of his/her capabilities to bring about desired outcomes of student engagement and learning' (Tschannen-Moran & Hoy, 2001, p. 783). As positive correlations between teacher efficacy and desirable outcomes on both the teacher level (e.g. enthusiasm and commitment) and the student level (e.g. achievement and motivation) have been established (e.g., Tschannen-Moran & Hoy, 2001), we compared average pretest and posttest scores on the teacher efficacy item from the teacher survey. The results are presented in Panel B of Table 8, showing that in the control and No

OTPD conditions, the average responses on the teacher efficacy item tend to decrease, while a small increase is observed in the OTPD condition. Additional t-tests demonstrated that the coefficients for the pretest to posttest change are significantly different between the two treatment conditions ($p \leq 0.01$). It is plausible that the treated teachers felt more capable of providing their students with additional support and instructions.

Third, given that attitudes towards technology influence its usage and acceptance, and that the effectiveness of OTPD initiatives, in turn, relies on the acceptance of online professional learning (Kao & Tsai, 2009; Smith & Sivo, 2012), we examined teachers' initial attitudes towards online professional learning. Therefore, the positive impact of the initiative could have resulted from more positive attitudes towards OTPD among those teachers who participated in the module. In both the pretest and posttest, teachers were asked to report on their attitudes towards online professional learning. Panel C of Table 8 provides an overview of the descriptive statistics from both tests for each condition and shows that attitudes towards OTPD do not differ between conditions, nor is there a significant change between the pretest and posttest in any of the conditions. We can therefore conclude that the results were not biased by differences in attitudes towards OTPD.

The final potential mechanism we identified is related to teachers' instructional practices. This concerns the transfer of teacher quality to actual teaching behaviour in the classroom (Desimone, 2009). To measure this, 20% of the schools in the No OTPD and OTPD conditions were visited for a classroom observation. Ultimately, we found no systematic differences between the treatment conditions for aspects such as the amount or type of help provided to the students while they went through the material. However, the duration of the introduction that preceded the educational game differed between the two intervention groups. While teachers in the No OTPD condition spent between 5 and 15 minutes introducing the topic, teachers in the OTPD condition consistently spent more than 15 minutes. Teachers' explanations about, for example, the purpose of the programme, the related practicalities, or a first introduction to the concepts covered, might have resulted in students having higher levels of motivation or prior knowledge before going through the material.

A limitation is that the number of teachers in the final sample was relatively limited. As a result, the findings concerning the underlying mechanisms that may explain the impact of OTPD need to be interpreted with caution. Furthermore, this implied that the power was insufficient to perform a causal mediation analysis on a larger share of the teacher characteristics that were measured in the teacher pretest (e.g. years of teaching experience and teaching economics). This type of analysis, as well as a higher number of classroom observations shedding light on changes in teaching behaviour, would have provided more insight into the mechanisms playing a role in the process from OTPD participation to student learning.

< Table 8 >

6. Discussion

It has been posited that the effectiveness of financial education is partly determined by the quality of teachers. Nevertheless, the evidence on the role of teacher professional development in the effectiveness of financial education as obtained by large-scale RCTs is limited. Therefore, the present study examined the extent to which teacher participation in an OTPD module enhanced students' financial literacy. This module was characterised by its scalability and the limited time investment required from teachers, which are desirable TPD features in times of curriculum reform. Two randomised controlled trials were conducted, with a total of 1827 students, 53 teachers and 47 schools participating.

The results demonstrated that the financial education programme increased students' financial knowledge (0.17 *SD)*. This finding is in line with existing literature, as Kaiser and Menkhoff (2020) found a similar average positive effect of 0.15 *SD* on financial knowledge in RCTs evaluating school-based financial education.[10] Our study also confirmed the findings of Batty et al. (2015), which demonstrated that even programmes of short duration (i.e. four or five lessons) are able to enhance students' financial knowledge. While it may be expected that programmes of sustained duration have the most impact, Kaiser and Menkhoff (2020) provided evidence supporting the hypothesis of Miller et al. (2015) by revealing declining marginal returns of programme intensity.[11] Specifically, an average effect size of 0.17 *SD* was shown for interventions lasting up to 5 hours. Potential explanations for the effectiveness of short programmes may be related to the 'novelty' of the material, and the opportunity to keep the environment well controlled (Cheung & Slavin, 2013). Our results further showed that providing teachers with access to the OTPD module did not significantly enhance the effectiveness of the programme on knowledge scores.

Focusing on the behaviour scores, no significant increase was found to result from following the programme itself. This contrasts with the findings of two recent meta-analyses, which revealed a significant impact of financial education interventions on financial behaviours; although the effect size was considerably smaller than for the impact on financial knowledge (Kaiser et al., 2020; Kaiser & Menkhoff, 2020). Nevertheless, our results showed an increase in behaviour scores (0.39 *SD* relative to the control condition) when the teachers were given the opportunity to participate in the OTPD

[10] In a meta-analysis of RCTs evaluating interventions in various educational domains (e.g. language and mathematics), Lortie-Forgues and Inglis (2019) found an average effect size of 0.06 *SD*. This implies that financial education interventions generally produce sizeable effects compared to interventions in other domains.

[11] Decreasing marginal returns of the duration of interventions have also been revealed in other educational domains (de Boer et al., 2014).

initiative. Also the difference between the two treatment conditions was shown to be significant. These results suggest that professional development initiatives that require relatively little time investment of teachers, are still effective in enhancing student learning. This contrasts the TPD model of Desimone (2009), in which sustained duration of initiatives is considered a key requirement for effective TPD initiatives.

Although effect sizes of a quarter of a standard deviation have been demonstrated for financial knowledge (Bruhn et al., 2016), the magnitude of the effect obtained in the present study appears large for behavioural items. One potential explanation is related to how financial behaviour was captured. Studies usually evaluate the impact on current behaviours of students, such as saving or intertemporal decision-making (Kalmi, 2018; Lührmann et al., 2018). Quasi-experiments tend to investigate long-term effects, such as the likelihood of debt during adolescence, student loan repayments or payday borrowing (e.g., Brown et al., 2016; Mangrum, 2019). However, as parents may still largely determine how children spend their (pocket) money at the age of 14-15, one could ask how informative reported changes in current behaviours are for the age group considered in our analysis. Ultimately, students of this age are unlikely to engage in a lot of financial decision-making. Therefore, we assessed whether students could detect and properly respond to receiving fraudulent bank emails, and if they were aware of how they should handle a bankcard. This implies that, strictly speaking, we measured students' intentions to engage in wise financial behaviours. Thus, the fact that the construct was measured differently in the present study, may provide an explanation for why the effect size that we found was larger in magnitude than those found in previous RCTs or quasi-experiments.

Combining the knowledge and behaviour scores to measure students' financial literacy led to the conclusion that the programme itself enhanced financial literacy (0.20 *SD*) and that access to the OTPD module enhanced this effectiveness (with an estimate of 0.43 *SD* relative to the control condition). These findings are in line with Compen et al. (2020), where teacher participation in a webinar series was shown to benefit student learning outcomes by 0.39 *SD*. The results for the separate knowledge and behaviour scores in the present study showed that the beneficial effect of OTPD was driven by increased performance on items measuring financial behaviour rather than items measuring financial knowledge. A plausible explanation may be that financial literacy is one of the few topics taught in schools that actively attempts to influence the behaviour of students, in addition to increasing their knowledge. While teachers in both treatment conditions might have been able to increase students' financial knowledge, the OTPD module might have been particularly effective in preparing teachers to influence students' behavioural intentions.

Our results have confirmed previous studies in demonstrating that financial education is effective in increasing students' financial literacy. In addition, we showed that students' behavioural

intentions benefit from their teachers engaging in professional development. Given that financial literacy has been shown to be a competence that persists throughout people's lives (Lusardi et al., 2010), these results are promising. Targeting today's youth may improve their financial decision-making at later ages, when they become more active consumers. Especially considering that the OTPD initiative only required approximately three hours of the teachers' time, and that the initiative is highly scalable, we recommend that policymakers enhance financial education programmes by encouraging teacher training initiatives.

There are a few limitations of this study which should be mentioned. First, the results are based on posttests that were completed immediately after the programme. The meta-analyses of Fernandes et al. (2014) and Kaiser and Menkhoff (2020) found that the effect sizes of financial education programmes are generally largest immediately after programme implementation and tend to decrease over time. Additionally, it has been shown that TPD effects tend to fade over time (Wolf & Peele, 2019). Therefore, it is likely that the effect sizes of the programme and OTPD module obtained in this chapter will also decrease. However, we were unable to measure the long-term effects due to the high attrition rate. Although Compen et al. (2020) demonstrated a positive impact of the evaluated TPD initiative on students' financial literacy both immediately after programme implementation, as well as six weeks later, an avenue for future research would be to estimate the long-term impact of OTPD. Finally, the limited number of participating teachers does not allow us to perform a causal analysis of how teacher characteristics can explain the positive impact of OTPD on student' test scores. In response to both limitations, future studies could consider providing schools with incentives to reduce the attrition rates in both the student and the teacher samples. The resulting increase in power would allow analysing longitudinal effects and gaining more understanding of the teacher-level mechanisms.

References

Agnew, S., & Cameron-Agnew, T. (2015). The influence of consumer socialisation in the home on gender differences in financial literacy. *International Journal of Consumer Studies, 39*(6), 630–638. https://doi.org/10.1111/ijcs.12179

Anderson. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association, 103*(484), 1481–1495. https://doi.org/10.1198/016214508000000841

Angrist, J., & Lavy, V. (2009). The effects of high stakes high school achievement awards: Evidence from a randomized trial. *American Economic Review, 99*(4), 1384–1414. https://doi.org/10.1257/aer.99.4.1384

Antonietti, A., Borsetto, A., & Iannello, P. (2016). A metacognitive approach to financial literacy. In C. Aprea, E. Wuttke, K. Breuer, N. K. Koh, P. Davies, B. Greimel-Fuhrmann, & J. S. Lopus (Eds.), *International Handbook of Financial Literacy* (pp. 57–68). Springer. https://doi.org/10.1007/978-981-10-0360-8

Bargagli-Stoffi, F., J., De Beckker, K., Maldonado, J. E., & De Witte, K. (2021). *Assessing sensitivity of machine learning predictions. A novel toolbox with an application to financial literacy* (arXiv:2102.04382)[Preprint]. https://arxiv.org/pdf/2102.04382.pdf

Batty, M., Collins, J. M., & Odders-White, E. (2015). Experimental evidence on the effects of financial education on elementary school students' knowledge, behavior, and attitudes. *The Journal of Consumer Affairs, 49*(1), 69–96. https://doi.org/10.1111/joca.12058

Benjamini, Krieger, & Yekutieli. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika, 93*(3), 491–507. https://doi.org/10.1093/biomet/93.3.491

Blue, L., Grootenboer, P., & Brimble, M. (2014). Financial literacy education in the curriculum: Making the grade or missing the mark? *International Review of Economics Education, 16*(Part A), 51–62. https://doi.org/10.1016/j.iree.2014.07.005

Brown, M., Grigsby, J., van der Klaauw, W., Wen, J., & Zafar, B. (2016). Financial education and the debt behavior of the young. *The Review of Financial Studies, 29*(9), 2490–2522. https://doi.org/10.1093/rfs/hhw006

Bruhn, M., de Souza Leão, L., Legovini, A., Marchetti, R., & Zia, B. (2016). The impact of high school financial education: Evidence from a large-scale evaluation in Brazil. *American Economic Journal: Applied Economics, 8*(4), 256–295. https://doi.org/10.1257/app.20150149

Chen, Y., Chen, N., & Tsai, C. (2009). The use of online synchronous discussion for web-based professional development for teachers. *Computers & Education, 53*(4), 1155-1166. https://doi.org/10.1016/j.compedu.2009.05.026

Cheung, A. C. K., & Slavin, R. E. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational Research Review, 9*, 88-113. https://doi.org/10.1016/j.edurev.2013.01.001

Compen, B., De Witte, K., & Schelfhout, W. (2018). The role of teacher professional development in financial literacy education: A systematic literature review. *Educational Research Review, 28*, 16-31. https://doi.org/10.1016/j.edurev.2018.12.001

Compen, B., De Witte, K., & Schelfhout, W. (2020). The impact of teacher engagement in an interactive webinar series on the effectiveness of financial literacy education. *British Journal of Educational Technology, 52*(1), 411-425. https://doi.org/10.1111/bjet.13013

Compen, B., & Schelfhout, W. (2021). Collaborative curriculum design in the context of fnancial literacy education. *Journal of Risk and Financial Management, 14*(6). https://doi.org/https://doi.org/10.3390/jrfm14060234

De Beckker, K., Compen, B., De Bock, D., & Schelfhout, W. (2019a). The capabilities of secondary school teachers to provide financial education. *Citizenship, Social and Economics Education, 18*(2), 66–81. https://doi.org/10.1177/2047173419850152

De Beckker, K., De Witte, K., & Van Campenhout, G. (2019b). Identifying financially illiterate groups: An international comparison. *International Journal of Consumer Studies 43*(5), 490-501. https://doi.org/10.1111/ijcs.12534

de Boer, H., Donker, A. S., & van der Werf, M. P. C. (2014). Effects of the attributes of educational interventions on students' academic performance: A meta-analysis. *Review of Educational Research, 84*(4), 509-545. https://doi.org/10.3102/0034654314540006

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher, 38*(3), 181–199. https://doi.org/10.3102/0013189X08331140

Fernandes, D., Lynch Jr., J. G., & Netemeyer, R. G. (2014). Financial literacy, financial education, and downstream financial behaviors. *Management Science, 60*(8), 1861–1883. https://doi.org/10.1287/mnsc.2013.1849

Fryer Jr., R. G. (2016). Information, non-financial incentives, and student achievement: Evidence from a text messaging experiment. *Journal of Public Economics, 144*, 109–121. https://doi.org/10.1016/j.jpubeco.2016.10.009

Harter, C. L., & Harter, J. F. R. (2012). Does a graduate course in personal finance for teachers lead to higher student financial literacy than a teacher workshop? *Journal of Consumer Education, 29*, 35-46. http://www.cefe.illinois.edu/JCE/archives/vol29.html

Hastings, J. S., Madrian, B. C., & Skimmyhorn, W. L. (2013). Financial literacy, financial education and economic outcomes. *Annual Review of Economics, 5*, 347-373. https://doi.org/10.1146/annurev-economics-082312-125807

Huston, S. J. (2010). Measuring financial literacy. *The Journal of Consumer Affairs, 44*(2), 296-316. https://doi.org/10.1111/j.1745-6606.2010.01170.x

Iacus, S. M., King, G., & Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Political Analysis, 20*(1), 1-24. https://doi.org/10.1093/pan/mpr013

Iterbeke, K., De Witte, K., Declercq, K., & Schelfhout, W. (2020a). The effect of ability matching and differentiated instruction in financial literacy education. Evidence from two randomised control trials. *Economics of Education Review, 78*. https://doi.org/10.1016/j.econedurev.2019.101949

Iterbeke, K., De Witte, K., & Schelfhout, W. (2020b). *The effects of computer-assisted adaptive instruction and elaborated feedback on learning outcomes. A randomized control trial* [In Press]. https://doi.org/10.1016/j.chb.2020.106666

Jones, D., Molitor, D., & Reif, J. (2019). What do workplace wellness programs do? Evidence from the Illinois Workplace Wellness Study. *The Quarterly Journal of Economics, 134*(4), 1747–1791. https://doi.org/10.1093/qje/qjz023

Kaiser, T., Lusardi, A., Menkhoff, L., & Urban, C. J. (2020). *Financial education affects financial knowledge and downstream behaviors*. https://doi.org/10.3386/w27057

Kaiser, T., & Menkhoff, L. (2020). Financial education in schools: A meta-analysis of experimental studies. *Economics of Education Review, 78*. https://doi.org/10.1016/j.econedurev.2019.101930

Kalmi, P. (2018). The effects of financial education: Evidence from Finnish lower secondary schools. *Economic Notes 47*(2-3). https://doi.org/doi.org/10.1111/ecno.12114

Kao, C. P., & Tsai, C. C. (2009). Teachers' attitudes toward web-based professional development, with relation to internet self-efficacy and beliefs about web-based learning. *Computers & Education, 53*(1), 66–73. https://doi.org/10.1016/j.compedu.2008.12.019

Klapper, L., Lusardi, A., & van Oudheusden, P. (2015). *Financial literacy around the world: Insights from the Standard & Poor's Ratings Services Global Financial Literacy Survey* [Report]. International Bank for Reconstruction and Development/World Bank. https://gflec.org/initiatives/sp-global-finlit-survey/

Lee, D. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies, 76*(3), 1071–1102. https://doi.org/10.1111/j.1467-937X.2009.00536.x

Lortie-Forgues, H., & Inglis, M. (2019). Rigorous large-scale educational RCTs are often uninformative: Should we be concerned? *Educational Researcher, 48*(3), 158–166. https://doi.org/10.3102/0013189X19832850

Lührmann, M., Serra-Garcia, M., & Winter, J. (2015). Teaching teenagers in finance: Does it work? *Journal of Banking & Finance, 54*, 160–174. https://doi.org/10.1016/j.jbankfin.2014.11.009

Lührmann, M., Serra-Garcia, M., & Winter, J. (2018). The impact of financial education on adolescents' intertemporal choices. *American Economic Journal: Economic Policy 10*(3), 309–332. https://doi.org/10.1257/pol.20170012

Lusardi, A., & Mitchell, O. S. (2007). Financial literacy and retirement preparedness: Evidence and implications for financial education. *Business Economics, 42*(1), 35-44. https://doi.org/10.2145/20070104

Lusardi, A., Mitchell, O. S., & Curto, V. (2010). Financial literacy among the young. *Journal of Consumer Affairs, 44*(2), 358–380. https://doi.org/10.1111/j.1745-6606.2010.01173.x

Lusardi, A., & Tufano, P. (2015). Debt literacy, financial experiences, and overindebtedness. *Journal of Pension Economics & Finance, 14*(4), 332–368. https://doi.org/doi:10.1017/S1474747215000232

Maldonado, J. E., & De Witte, K. (2021). The impact of information provision to parents: Experimental evidence on student outcomes. *Journal of Behavioral and Experimental Finance, 31*, 1-10. https://doi.org/https://doi.org/10.1016/j.jbef.2021.100548

Maldonado, J. E., De Witte, K., & Declercq, K. (2021). The effects of parental involvement in homework: two randomised controlled trials in financial education. *Empirical Economics*. https://doi.org/10.1007/s00181-021-02058-8

Mangrum, D. (2019). *Personal finance education mandates and student loan repayment* [Job Market Paper]. https://doi.org/10.2139/ssrn.3349384

Matzat, U. (2013). Do blended virtual learning communities enhance teachers' professional development more than purely virtual ones? A large scale empirical comparison. *Computers & Education, 60*(1), 40-51. https://doi.org/10.1016/j.compedu.2012.08.006

Merchie, E., Tuytens, M., Devos, G., & Vanderlinde, R. (2016). Evaluating teachers' professional development initiatives: towards an extended evaluative framework. *Research Papers in Education, 13*(2), 143-168. https://doi.org/10.1080/02671522.2016.1271003

Miller, M., Reichelstein, J., Salas, C., & Zia, B. (2015). Can you help someone become financially capable? A meta-analysis of the literature. *The World Bank Research Observer, 30*(2), 220–246. https://doi.org/10.1093/wbro/lkv009

Nicolini, G. (2019). *Financial literacy in Europe. Assessment methodologies and evidence from European countries*. Routledge.

OECD. (2016). *Financial education in Europe: Trends and recent developments*. OECD Publishing. https://doi.org/10.1787/9789264254855-en

OECD. (2017). *PISA 2015 results (volume IV): Students' financial literacy*. OECD Publishing. https://doi.org/10.1787/9789264270282-en

Pak, T. Y. (2018). Financial literacy and high-cost borrowing: Exploring the mechanism. *International Journal of Consumer Studies, 42*(3), 283-294. https://doi.org/10.1111/ijcs.12429

Partnership for 21st Century Learning. (2019). *Framework for 21st century learning definitions*. https://www.battelleforkids.org/networks/p21/frameworks-resources

Riitsalu, L., & Põder, K. (2016). A glimpse of the complexity of factors that influence financial literacy. *International Journal of Consumer Studies, 40*(6), 722-731. https://doi.org/10.1111/ijcs.12291

Sevim, N., Temizel, F., & Sayılır, Ö. (2012). The effects of financial literacy on the borrowing behaviour of Turkish financial consumers. *International Journal of Consumer Studies, 36*(5), 573-579. https://doi.org/10.1111/j.1470-6431.2012.01123.x

Smith, J. A., & Sivo, S. A. (2012). Predicting continued use of online teacher professional development and the influence of social presence and sociability. *British Journal of Educational Technology 43*(6), 871-882. https://doi.org/10.1111/j.1467-8535.2011.01223.x

Swinton, J., DeBerry, T., Scafidi, B., & Woodard, H. (2007). The impact of financial education workshops for teachers on students' economic achievement. *The Journal of Consumer Education, 24*, 63–77.

Tauchmann, H. (2014). Lee (2009) treatment-effect bounds for nonrandom sample selection. *The Stata Journal, 14*(4), 884–894. https://doi.org/10.1177/1536867X1401400411

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*(7), 783-805. https://doi.org/10.1016/S0742-051X(01)00036-1

Urban, C., Schmeiser, M., Collins, J. M., & Brown, A. (2020). The effects of high school personal financial education policies on financial behavior. *Economics of Education Review, 78*. https://doi.org/10.1016/j.econedurev.2018.03.006

Van Campenhout, G., De Witte, K., & De Beckker, K. (2017). Financiele vorming op school. In *De geslaagde school* (pp. 153 - 181).

van Rooij, M., Lusardi, A., & Alessie, R. (2011). Financial literacy and stock market participation. *Journal of Financial Economics, 101*(2), 449-472. https://doi.org/10.1016/j.jfineco.2011.03.006

van Rooij, M., Lusardi, A., & Alessie, R. (2012). Financial literacy, retirement planning and household wealth. *The Economic Journal, 122*(560), 449-478. https://doi.org/10.1111/j.1468-0297.2012.02501.x.

Walstad, W. B., Rebeck, K., & MacDonald, R. A. (2010). The effects of financial education on the financial knowledge of high school students. *The Journal of Consumer Affairs, 44*(2), 336-357. https://doi.org/10.1111/j.1745-6606.2010.01172.x

Way, W. L., & Holden, K. (2010). Teachers' background and capacity to teach personal finance: Results of a national study [Report]. *Journal of Financial Counseling and Planning, 20*(2), 64-78.

World Economic Forum. (2015). *The skills needed in the 21st century*. https://widgets.weforum.org/nve-2015/chapter1.html

Tables

Table 1: Research design and timing

Panel A: Experiment 1				
	Pretest	**OTPD module**	**Financial literacy education**	**Posttest**
	January 2018	*February – March 2018*	*February – March 2018*	*February – March 2018*
Control condition	✓	-	-	✓
No OTPD condition	✓	-	✓	✓
OTPD condition	✓	✓	✓	✓
Panel B: Experiment 2				
	Pretest	**OTPD module**	**Financial literacy education**	**Posttest**
	February – March 2018	*April 2018*	*April 2018*	*April 2018*
Control condition	✓	-	-	✓
No OTPD condition	✓	-	✓	✓
OTPD condition	✓	✓	✓	✓

Table 2: Descriptive statistics at the student, teacher and school levels

Panel A: Students			
	Control	**No OTPD**	**OTPD**
Background characteristics			
Gender (female)	0.534	0.506	0.501
	(0.055)	(0.056)	(0.057)
Age	14.26	14.48	14.25
	(0.138)	(0.136)	(0.147)
SES indicator	1.945	1.870	2.078
	(0.079)	(0.082)	(0.080)
Language (no Dutch)	0.142	0.358**	0.143
	(0.069)	(0.068)	(0.042)
Motivation	4.172	4.057*	4.219
	(0.039)	(0.044)	(0.031)
Self-efficacy	3.499	3.431	3.425
	(0.050)	(0.056)	(0.064)
Financial literacy - pretest			
Knowledge score	1.865	1.854	2.042
	(0.095)	(0.098)	(0.104)
Behaviour score	1.159	1.035	1.256
	(0.086)	(0.087)	(0.077)
Total score	3.023	2.888	3.298
	(0.164)	(0.166)	(0.167)
Financial literacy - posttest			
Knowledge score	1.962	2.106	2.371**
	(0.110)	(0.125)	(0.122)
Behaviour score	1.044	1.075	1.500***
	(0.077)	(0.080)	(0.076)
Total score	3.007	3.181	3.822***
	(0.187)	(0.189)	(0.177)
N	768	492	567
Joint balance test	.000		
Panel B: Teachers			
	Control	**No OTPD**	**OTPD**
Financial literacy - pretest			
Total score	5.708	4.438**	4.769*
	(0.322)	(0.429)	(0.407)
Financial literacy - posttest			
Total score	5.875	5.688	6.077
	(0.300)	(0.387)	(0.380)
Participation in OTPD			
Login at least once	N.A.	N.A.	76.9%
Number of clicks	N.A.	N.A.	157
N	24	16	13
Joint balance test	.003		
Panel C: Schools			
	Control	**No OTPD**	**OTPD**
Public school	0.353	0.313	0.357
	(0.118)	(0.122)	(0.132)
Academic school	0.588	0.313	0.500
	(0.120)	(0.123)	(0.137)
N	17	16	14
Joint balance test	.000		

Note: Only students and teachers that completed both the pretest and the posttest are included. Gender is a dummy variable taking the value 1 if the student is female. Age reflects the age at pretest. SES indicator is measured using the number of holidays abroad as a proxy (0 = none, 1 = one, 2 = two, 3 = more than two). Language is a dummy variable taking the value 1 if a foreign language is spoken at home. Motivation is the response to the statement: 'I find it important to have knowledge on financial matters, such as saving, costs, and payment methods' (answered on a Likert scale from 0 = completely disagree to 5 = completely agree). Self-efficacy is the response to the statement: 'I find that I have sufficient knowledge on financial matters, such as saving, costs, and payment methods' (answered on a Likert scale from 0 = completely disagree to 5 = completely agree). Pretest and posttest scores are unstandardised and scored on a scale of 9 (total score), 6 (knowledge score), and 3 (behaviour score). Login at least once reflects the share of teachers in the OTPD condition who logged in into the OTPD module. Number of clicks refers to the average number of clicks of the teachers who logged in at least once. Public school is a dummy variable taking the value of 1 if the school is public rather than private, and academic school is a dummy variable taking the value of 1 if the school solely offered the academic track. The joint balance test rows indicate the *p*-values from jointly testing whether the variables at the student, teacher and school levels, respectively, are jointly insignificant in predicting enrolment into the treatment conditions. Standard errors clustered at the school level in parentheses. Significance levels correspond to differences relative to the control condition, derived from t-tests. *p < .10, **p < .05, ***p < .01.

Table 3: Differential attrition

No OTPD	0.196**	0.173*
	(0.084)	(0.089)
OTPD	0.137**	0.130*
	(0.059)	(0.066)
Pretest score	-0.030**	-0.021
Controls	No	Yes
N	2496	2496
No OTPD = OTPD	.479	.632

Note: N refers to the number of students completing the pretest and for whom there are no missing values. The control condition serves as the reference group. Control variables are standardised pretest score, gender, SES indicator, language, motivation, self-efficacy, pretest score teacher (standardised), and type of school (public and academic). In the combined sample, grade is additionally controlled for. The *p*-values resulting from *F*-tests (No OTPD = OTPD) comparing the effects of the treatment conditions are represented in the bottom row. Standard errors clustered at the school level in parentheses Standard errors clustered at the school level in parentheses. * $p \leq .10$ ** $p \leq .05$ *** $p \leq .01$.

Table 4: Empirical results: ITT

	Knowledge score		**Behaviour score**		**Total score**	
No OTPD	0.135	0.172*	0.056	0.131	0.129	0.197*
	(0.109)	(0.100)	(0.091)	(0.098)	(0.115)	(0.108)
OTPD	0.295***	0.294***	0.383***	0.386***	0.424***	0.425***
	(0.102)	(0.096)	(0.073)	(0.078)	(0.095)	(0.092)
Pretest score	0.214***	0.192***	0.271***	0.247***	0.304***	0.275***
	(0.017)	(0.016)	(0.025)	(0.027)	(0.019)	(0.020)
Controls	No	Yes	No	Yes	No	Yes
R^2	.065	.075	.110	.127	.134	.150
N	1827	1827	1827	1827	1827	1827
No OTPD = OTPD	.067	.222	.001	.005	.001	.020

Note: Dependent variables represent standardised posttest scores. The control condition serves as the reference group. Control variables are standardised pretest score, gender, grade, SES indicator, language, motivation, self-efficacy, pretest score teacher (standardised), public, and general education. The *p*-values resulting from *F*-tests (No OTPD = OTPD) comparing the effects of the treatment conditions are represented in the bottom row. Standard errors clustered at the school level in parentheses. * $p \leq .10$ ** $p \leq .05$ *** $p \leq .01$.

Table 5: Empirical results: 2SLS

Panel A: Having logged in at least once as the indicator for participation in OTPD						
First stage						
	Knowledge score		**Behaviour score**		**Total score**	
Instrument (OTPD)	0.563***	0.644***	0.563***	0.644***	0.563***	0.644***
	(0.135)	(0.114)	(0.135)	(0.114)	(0.135)	(0.114)
Controls	No	Yes	No	Yes	No	Yes
Second stage						
	Knowledge score		**Behaviour score**		**Total score**	
Treatment	0.136	0.191**	0.057	0.171*	0.130	0.232**
	(0.108)	(0.093)	(0.089)	(0.093)	(0.113)	(0.101)
Login	0.284	0.189	0.580***	0.394***	0.525**	0.354**
	(0.180)	(0.157)	(0.196)	(0.139)	(0.209)	(0.157)
Pretest score	0.220***	0.194***	0.283***	0.252***	0.315***	0.280***
	(0.018)	(0.017)	(0.024)	(0.026)	(0.020)	(0.021)
Controls	No	Yes	No	Yes	No	Yes
R^2	.048	.067	.083	.117	.098	.133
N	1827	1827	1827	1827	1827	1827
Panel B: Number of clicks as the indicator for participation in OTPD						
First stage						
	Knowledge score		**Behaviour score**		**Total score**	
Instrument (OTPD)	0.793***	0.958***	0.793***	0.958***	0.793***	0.958***
	(0.252)	(0.237)	(0.252)	(0.237)	(0.252)	(0.237)
Controls	No	Yes	No	Yes	No	Yes
Second stage						
	Knowledge score		**Behaviour score**		**Total score**	
Treatment	0.135	0.195**	0.056	0.180**	0.128	0.240**
	(0.108)	(0.091)	(0.090)	(0.090)	(0.114)	(0.099)
Clicks	0.202*	0.127	0.412***	0.265***	0.372**	0.238**
	(0.122)	(0.098)	(0.159)	(0.100)	(0.151)	(0.098)
Pretest score	0.211***	0.189**	0.264***	0.241***	0.298***	0.269***
	(0.021)	(0.018)	(0.031)	(0.028)	(0.027)	(0.023)
Controls	No	Yes	No	Yes	No	Yes
R^2	.045	.067	.044	.104	.074	.128
N	1827	1827	1827	1827	1827	1827

Note: Dependent variables represent standardised posttest scores. The control condition serves as the reference group. Treatment is a dummy variable equal to 1 if the student is allocated to one of the two treatment conditions, and therefore represents the average effect of the financial education programme on student achievement for the no OTPD and OTPD conditions combined. Control variables are standardised pretest score, gender, SES indicator, language, motivation, self-efficacy, grade, pretest score teacher (standardised), public, and general education. Standard errors clustered at school level in parentheses. * $p \leq .10$ ** $p \leq .05$ *** $p \leq .01$.

Table 6: Robustness tests

Panel A: ITT – School level

	Knowledge score		Behaviour score		Total score	
No OTPD	0.269**	0.308**	0.175**	0.191**	0.291**	0.328***
	(0.127)	(0.130)	(0.086)	(0.084)	(0.121)	(0.119)
OTPD	0.286**	0.306***	0.278***	0.280***	0.359***	0.376***
	(0.108)	(0.113)	(0.092)	(0.098)	(0.104)	(0.111)
Pretest score	0.373***	0.333***	0.510***	0.512***	0.551***	0.523***
	(0.117)	(0.124)	(0.074)	(0.080)	(0.110)	(0.120)
Controls	No	Yes	No	Yes	No	Yes
R^2	.379	.398	.621	.640	.571	.589
N	47	47	47	47	47	47
No OTPD = OTPD	.894	.988	.296	.366	.562	.696

Panel B: Difference-in-differences

	Knowledge score		Behaviour score		Total score	
No OTPD	0.126	0.126	0.186	0.186	0.189*	0.189*
	(0.097)	(0.097)	(0.113)	(0.113)	(0.107)	(0.107)
OTPD	0.179*	0.180*	0.311***	0.311***	0.301***	0.302***
	(0.098)	(0.098)	(0.074)	(0.074)	(0.094)	(0.095)
Controls	No	Yes	No	Yes	No	Yes
R^2	.013	.048	.024	.058	.027	.076
N	1827	1827	1827	1827	1827	1827
No OTPD = OTPD	.571	.571	.273	.274	.218	.218

Panel C: Coarsened Exact Matching

	Total score	***N***
No OTPD	0.119**	1017
	(0.053)	
OTPD	0.371***	1104
	(0.048)	
OTPD = No OTPD	0.214***	812
	(0.057)	

Panel D: Lee Bounds

	Total score	**Lower bound**	**Upper bound**	**95% CI of effect**	***N***
No OTPD	0.084*	-0.181***	0.417***	[-0.270, 0.504]	1260
	(0.047)	(0.054)	(0.053)		
OTPD	0.394***	0.190***	0.541***	[0.096, 0.630]	1135
	(0.044)	(0.057)	(0.054)		
OTPD = No OTPD	0.310***	0.237***	0.453***	[0.130, 0.554]	1059
	(0.050)	(0.065)	(0.061)		

Panel E: Multiple testing correction

	Knowledge score		Behaviour score		Total score	
No OTPD	0.135	0.172	0.056	0.131	0.129	0.197
	(0.109)	(0.100)	(0.091)	(0.098)	(0.115)	(0.108)
OTPD	0.295**	0.294**	0.383***	0.386***	0.424***	0.425***
	(0.102)	(0.096)	(0.073)	(0.078)	(0.095)	(0.092)

Pretest score	0.214***	0.192***	0.271***	0.247***	0.304***	0.275***
	(0.017)	(0.016)	(0.025)	(0.027)	(0.194)	(0.020)
Controls	No	Yes	No	Yes	No	Yes
R^2	.065	.075	.110	.127	.134	.150
N	1827	1827	1827	1827	1827	1827
No OTPD = OTPD	.095	.202	.004	.014	.004	.031

Panel F: Two-way clustering

	Knowledge score		**Behaviour score**		**Total score**	
No OTPD	0.135	0.172*	0.056	0.132	0.129	0.197*
	(0.108)	(0.099)	(0.093)	(0.101)	(0.114)	(0.108)
OTPD	0.295***	0.294***	0.383***	0.386***	0.424***	0.425***
	(0.099)	(0.093)	(0.072)	(0.077)	(0.092)	(0.089)
Pretest score	0.214***	0.192***	0.271***	0.247***	0.304***	0.275***
	(0.016)	(0.014)	(0.024)	0.025)	(0.017)	(0.018)
Controls	No	Yes	No	Yes	No	Yes
R^2	.065	.075	.110	.127	0.134	.150
N	1827	1827	1827	1827	1827	1827
No OTPD = OTPD	.074	.216	.000	.002	.001	.015

Note. Dependent variables in Panel A represent standardised posttest scores at the school level, whereas dependent variables in Panel B represent standardised differences between pretest and posttest scores at the student level. Dependent variables in Panels C, D, E and F represent standardised posttest scores at the student level. The control condition serves as the reference category in all panels. Control variables in Panels B, E and F are standardised pretest score, gender, grade, SES indicator, language, motivation, self-efficacy, pretest score teacher (standardised), public, and general education. In Panel C, the following variables were included in the matching procedure: pretest score student standardised (scoring either in the lower or upper half of the distribution), gender, SES proxy, language, importance of knowledge and self-assessed knowledge. The multivariate L1 distances for all three scores approached zero in the matched sample. In Panel D, the trimming proportions are 23.8% (No OTPD – Control), 15.9% (OTPD – Control) and 9.4% (No OTPD – OTPD). In Panel E, the indicated significance levels are based on q-values calculated as in Benjamini et al. (2006) and Anderson (2008) to correct for multiple hypothesis testing. The *p*-values resulting from *F*-tests (No OTPD = OTPD) comparing the effects of the treatment conditions are represented in the bottom row of each panel. Standard errors clustered at the school level in all panels, except in Panels C and D. * $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$.

Table 7: Heterogeneity analysis

	Pretest score		**Gender**		**Socioeconomic status**		**Language**	
	Low	*High*	*Male*	*Female*	*Low*	*High*	*Dutch*	*No Dutch*
No OTPD	0.134	0.201	0.054	0.270**	0.219**	0.154	0.101	0.373***
	(0.092)	(0.122)	(0.093)	(0.107)	(0.094)	(0.101)	(0.092)	(0.114)
OTPD	0.317***	0.396***	0.214***	0.505***	0.291***	0.387***	0.375***	0.290***
	(0.081)	(0.089)	(0.080)	(0.080)	(0.090)	(0.082)	(0.077)	(0.100)
Pretest score	0.256***	0.058**	0.265***	0.181***	0.270***	0.209***	0.242***	0.160***
	(0.043)	(0.058)	(0.028)	(0.025)	(0.026)	(0.023)	(0.020)	(0.041)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	.103	.069	.162	.156	.179	.139	0.146	0.158
N	1128	699	884	943	606	1221	1461	366
No OTPD = OTPD	.034	.009	.031	.025	.478	.010	.002	.398

Note. Dependent variables represent standardised total posttest scores. The control condition serves as the reference category. Pretest score is a binary variable coded 0 when the student obtained a standardised pretest score lower than zero, and 1 when a pretest score higher than zero was obtained. Socio-economic status is a binary variable coded 0 when the student went on less than two holidays abroad in the last year, and 1 when the student had been on at least two holidays abroad. Language is a binary variable coded 0 when the student speaks Dutch at home, and 1 when another language is spoken at home. The following background characteristics are controlled for: standardised pretest score, gender, grade, SES proxy, language, importance of knowledge, self-assessed knowledge, standardised pretest score (teacher), academic education and public school. The *p*-values resulting from *F*-tests (No OTPD = OTPD) comparing the effects of the treatment conditions are represented in the bottom row. Standard errors clustered at the school level in parentheses. * $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$.

Table 8: Teachers' test scores, teacher efficacy and attitudes towards OTPD

Panel A: Average total scores of teacher pretest and posttest			
	Control	**No OTPD**	**OTPD**
Pretest	5.708	4.438***	4.769*
	(0.259)	(0.376)	(0.441)
Posttest	5.875	5.688	6.077
	(0.250)	(0.384)	(0.366)
Difference	0.167	1.250*	1.308*
	(0.305)	(0.496)	(0.548)
N	24	16	13
Panel B: Average scores on teacher efficacy in pretest and posttest			
	Control	**No OTPD**	**OTPD**
Pretest	6.042	5.750	6.231
	(0.153)	(0.266)	(0.231)
Posttest	5.667	4.813**	6.308*
	(0.223)	(0.379)	(0.133)
Difference	-0.375	-0.938*	0.077*
	(0.168)	(0.309)	(0.178)
N	24	16	13
Panel C: Attitudes towards online teacher professional development			
	Control	**No OTPD**	**OTPD**
Pretest	5.292	4.625	5.000
	(0.229)	(0.397)	(0.340)
Posttest	5.333	4.733	5.077
	(0.253)	(0.384)	(0.309)
Difference	0.042	0.067	0.077
	(0.259)	(0.316)	(0.459)
N	24	16	13

Note: The average number of questions answered correctly on the 9-item tests (Panel A), the average response to the statement: 'I find that I have sufficient knowledge on payment methods to teach about this in the second/third year of secondary education' (answered on a Likert scale from 1 = completely disagree to 7 = completely agree) (Panel B), and the average response to the statement: 'I feel positive about individual, online learning on a digital platform for teachers' (answered on a Likert scale from 1 = completely disagree to 7 = completely agree) (Panel C). Standard errors in parentheses. Significance levels correspond to differences relative to the control condition, derived from t-tests. *p ≤ .10, **p ≤ .05, ***p ≤ 0.01.

Figures

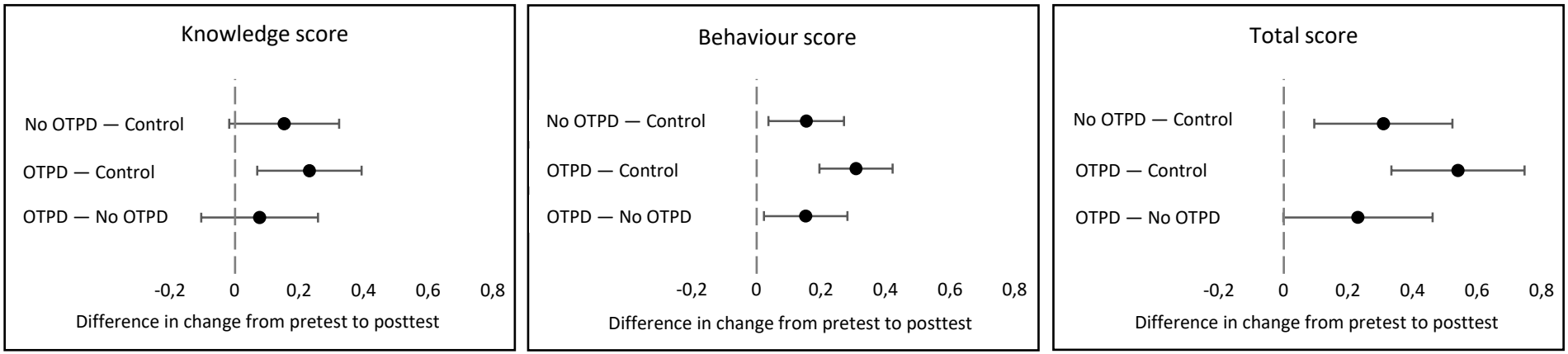

Figure 1. Comparison between conditions in average change from pretest to posttest.

Note: Dots represent the difference between conditions in score improvement from pretest to posttest and are surrounded by their 95% confidence intervals. Confidence intervals containing zero imply that the two conditions do not differ significantly in the average score change. Graphs are based on the combined sample (N = 1827).

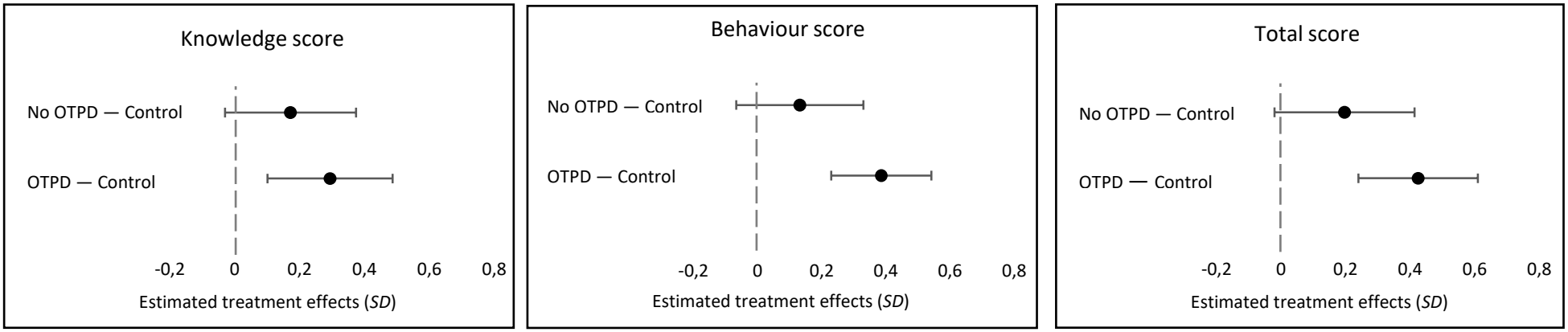

Figure 2. Comparison between conditions in the estimated treatment effects (ITT).

Note: Dots represent the difference in SD between conditions in estimated treatment effects and are surrounded by their 95% confidence intervals. Confidence intervals containing zero imply that the two conditions do not differ significantly in the estimated effects. Graphs are based on the combined sample (N = 1827).