Learning and Individual Differences 87 (2021) 101982

Contents lists available at ScienceDirect

Learning and Individual Differences

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

Why and for whom does personalizing math problems enhance performance? Testing the mediation of enjoyment and cognitive load at different ability levels

Eva Van de Weijer-Bergsma[a,*], Sanne H.G. Van der Ven[b]

[a] Utrecht University, Faculty of Social and Behavioral Sciences, Department of Pedagogical and Educational Sciences, P.O. Box 80140, 3508, TC, Utrecht, the Netherlands

[b] Radboud University, Behavioral Science Institute, Nijmegen, the Netherlands

ARTICLE INFO

Keywords:
Personalization
Mathematics
Working memory
Enjoyment
Cognitive load

ABSTRACT

In two context personalization studies, we examined (a) enjoyment and cognitive load as two potential mechanisms explaining the effects of context personalization on mathematical word problem performance, and (b) whether individual differences in math, reading and working memory ability moderated these effects. In both studies (Study 1: $N = 238$; Study 2: $N = 149$) primary school students from 6th grade completed math word problems in either a personalized condition or a control condition. Students rated their enjoyment and experienced cognitive load after each problem. Moderated mediation models showed that while ability, enjoyment and cognitive load significantly predicted performance, (a) personalization did not affect word problem performance, enjoyment or cognitive load, and (b) the three different abilities did not moderate these relations. The findings are discussed in light of three personalization principles (depth, grain size, ownership) and complexity in different steps of math problem solving.

1. Introduction

Mathematical proficiency is important for the prospects of both individuals (Parsons & Bynner, 2006) and society (Hanushek & Woessmann, 2010). Improving mathematics education is therefore of great importance. In mathematics education, word problems are often used, as they link problem solving abilities to students' real life experiences. Meaningful contexts can be helpful, but the extra text in word problems can also pose difficulties, especially when the context is unfamiliar or unappealing. Because the same context may not fit all students' experiences equally, personalizing the context to each student's interest can make the learning content meaningful, relevant and interest-driven (Walkington & Bernacki, 2018), and may improve mathematical learning (Bernacki & Walkington, 2014; Cordova & Lepper, 1996). However, the underlying mechanisms that make context personalization effective have not yet been examined systematically. Moreover, personalization may not be equally effective for all students: it may especially benefit lower-ability students (Davis-Dorsey et al., 1991; Ku & Sullivan, 2000; Walkington et al., 2013), while possibly even being counterproductive for better able students.

In the present paper, we investigated why and for whom personalization is effective. We propose two potential mechanisms of personalization: increased motivation or enjoyment and reduced cognitive load. Furthermore, we investigated whether the effectiveness varies with math, reading and working memory (WM) ability, and whether enjoyment and cognitive load explain these differential effects.

1.1. *The effects of personalization*

In context personalization studies, information in textbook instruction or math problems is replaced by individuals' and their friends' names, favorite places, objects, and activities. In most studies, context personalization is based on individual interests, but sometimes group interests are used. In some context personalization studies (Bates & Wiest, 2004; López & Sullivan, 1992), but more predominantly in multimedia research, a self-referencing approach is used, in which 'you' or the name of the learner is incorporated (Moreno & Mayer, 2000).

Both immediate performance and learning gains in mathematics

Abbreviation: WM, working memory.

* Corresponding author.
E-mail address: e.vandeweijer@uu.nl (E. Van de Weijer-Bergsma).

https://doi.org/10.1016/j.lindif.2021.101982
Received 20 May 2020; Received in revised form 4 February 2021; Accepted 8 February 2021
Available online 19 March 2021

have been shown to improve by personalizing the context of instruction (Akinsola & Awofala, 2009; Harter & Ku, 2008; López & Sullivan, 1992; Ross & Anand, 1987), word problem tests (Davis-Dorsey et al., 1991; Kosh, 2016), or both (Ku & Sullivan, 2002). Ku and Sullivan (2002), for example, conducted a study with 136 fourth-grade students and compared personalized and non-personalized instruction. In both conditions, students did better on personalized than on non-personalized problems, but students in the personalized instruction condition improved more from pretest to posttest than the control group. Likewise, self-referencing can have beneficial effects on performance on cognitive (learning) tasks (Moreno & Mayer, 2000; Sui & Humphreys, 2015).

1.2. Mechanisms of personalization

Two as yet untested mechanisms, which are not mutually exclusive, have been proposed for explaining the beneficial effects of personalization: (1) increased motivation and (2) reduced cognitive load (Ku & Sullivan, 2000).

The first proposed mechanism is enhanced motivation. Motivation can be defined as the willingness to exert effort to achieve a goal and thus consists of the affective component of willingness (often measured as interest and enjoyment), and the behavioral component of task effort and engagement. Motivation can range from intrinsic, when students undertake an activity because it sparks interest or satisfaction, to extrinsic, when a student undertakes an activity for reasons that lie beyond the activity itself (e.g., to gain a high grade). According to the Self-Determination Theory (Ryan & Deci, 2000), motivation is at its highest quality and most volitional when an individual experiences autonomy, competence and relatedness. Context personalization can be assumed to trigger intrinsic motivation: students are more likely to feel autonomous, competent and related to the context in which the problem is presented when this context is adapted to their individual interest. Context personalization can make students enjoy the task more and exert more effort (Walkington & Bernacki, 2014). As motivation involves several components, studies on context personalization and motivation often use different approaches depending on (the clarity of) the theoretical framework used. Most studies focused on examining the effects on two types of interest: situational interest and individual interest (Bernacki & Walkington, 2014; Høgheim & Reber, 2015, 2017). Situational interest refers to a state of heightened attention and engagement in a specific task. It is elicited by features of a task that trigger enjoyment or a sense of task value. Situational interest can lead to individual interest, which refers to more prolonged engagement and a continuing preference for a certain subject (e.g., mathematics) (Hidi & Renninger, 2006). Other studies included measures of self-efficacy (Akinsola & Awofala, 2009) or task effort (Høgheim & Reber, 2015, 2017). In some studies, however, the theoretical framework was less clear, which was reflected in the assessment of a mix of motivational constructs (e.g., interest, difficulty, enjoyment, familiarity) in a single measure (Ku & Sullivan, 2000, 2002; López & Sullivan, 1992). Together, the results from these studies indicate that mathematical context personalization can increase various motivational components. In the present study we focus on enjoyment as a reflection of intrinsic motivation.

The second proposed mechanism through which personalization can exert its effect is cognitive load reduction. In Cognitive Load Theory, cognitive load represents the amount of WM resources an activity requires from a person (Sweller, 1994). Since WM resources are limited, information processing can be hampered when cognitive load is (too) high. Context personalization may help to reduce this load; as problems are adapted to the interests of students, students are more familiar with the problem context. This helps them to bridge the gap between existing and new knowledge (Davis-Dorsey et al., 1991; Ross & Anand, 1987). A greater familiarity may ease cognitive load through activation of existing mental schemas, thus freeing up students' limited WM resources (Guida et al., 2013; Mayer et al., 2004). This assumption was confirmed by Guida, Tardieu & Nicolas (2009), who showed that WM performance, measured with a reading span task, was enhanced when the content of the material to be remembered was personalized.

1.3. Variability in the effectiveness of personalization

Despite these positive findings, several other studies failed to find beneficial effects of context personalization of instruction (Høgheim & Reber, 2015, 2017; McLaren et al., 2006) or word problem tests (Bates & Wiest, 2004; Cakir & Simsek, 2010). Such null results may be caused by ineffective personalization, in which the context is adapted in a shallow or irrelevant way. Indeed, Kosh (2016) found that in 8th grade students, performance on personalized word problems was only enhanced for personalized problem they rated as interesting.

Walkington and Bernacki (2018) proposed three interacting design principles for effective personalization: depth, grain size and ownership. With regard to depth, ideally not only the topics, but also the calculation strategies in the math problem are personalized to the interest. With regard to grain size, a fine-grained approach to personalization (e.g., basketball) may be more effective than a coarse-grained approach (e.g., sports). With regard to ownership, giving students more control should be more effective, for example by letting students actively choose content or tasks. A failure to find effects of personalization may thus be due to a lack of sufficient depth, grain size and/or ownership in the task.

Alternatively, the effect of personalization may also vary across students and even be detrimental in some. Notably, the effectiveness of personalization may depend on the student's ability (Bates & Wiest, 2004; Walkington et al., 2013). Indeed, some studies showed that context personalization had larger effects in students with low math ability (Harter & Ku, 2008; Ku & Sullivan, 2000; Walkington et al., 2013), while effects were even negative for high-math-ability students working on easier problems (Walkington et al., 2013). Personalization may then distract attention from the problem and become disruptive.

Besides math ability, reading ability may also moderate the effects of personalization. Students with low reading ability are poorer at mathematical problem solving (Hickendorff, 2013b; Vilenius-Tuohimaa et al., 2008) and may thus also benefit more from personalization. However, Bates and Wiest (2004) did not find an effect of math problem personalization at any reading ability level, but their reading measure was subjective (teacher report) and coarse (high, medium or low).

A third potential moderator is WM. We do not know of any studies investigating WM as a moderator of the effect of context personalization, but since it plays a crucial role in word problem solving (e.g., for a review, see Friso-van den Bos et al., 2013), students with low WM ability may benefit more, as personalization should reduce the cognitive load on their more limited WM resources. Initial evidence for this assumption comes from the study by Guida et al. (2009), who found that participants with low WM reading span scores benefited the most from personalization of the WM task.

1.4. Mechanisms to understand variability in the effectiveness of personalization

We have thus two proposed mechanisms through which personalization may work: increased enjoyment and reduced cognitive load. Furthermore, personalization may be especially effective for students of lower ability. This leads to our hypothesis that the expected mediational effects of increased motivation and reduced cognitive load are stronger in students of lower ability: a conditional process model in Hayes' (2018) terminology, sometimes also called moderated mediation. Fig. 1 presents all hypothesized effects with their expected sign (positive or negative). Fig. 1A shows the conceptual model, while Fig. 1B shows the corresponding statistical model, in which moderation effects are operationalized as interaction terms. We expect personalization to enhance performance, especially so or perhaps only in children of lower ability. We thus expect negative moderation of the direct effect of

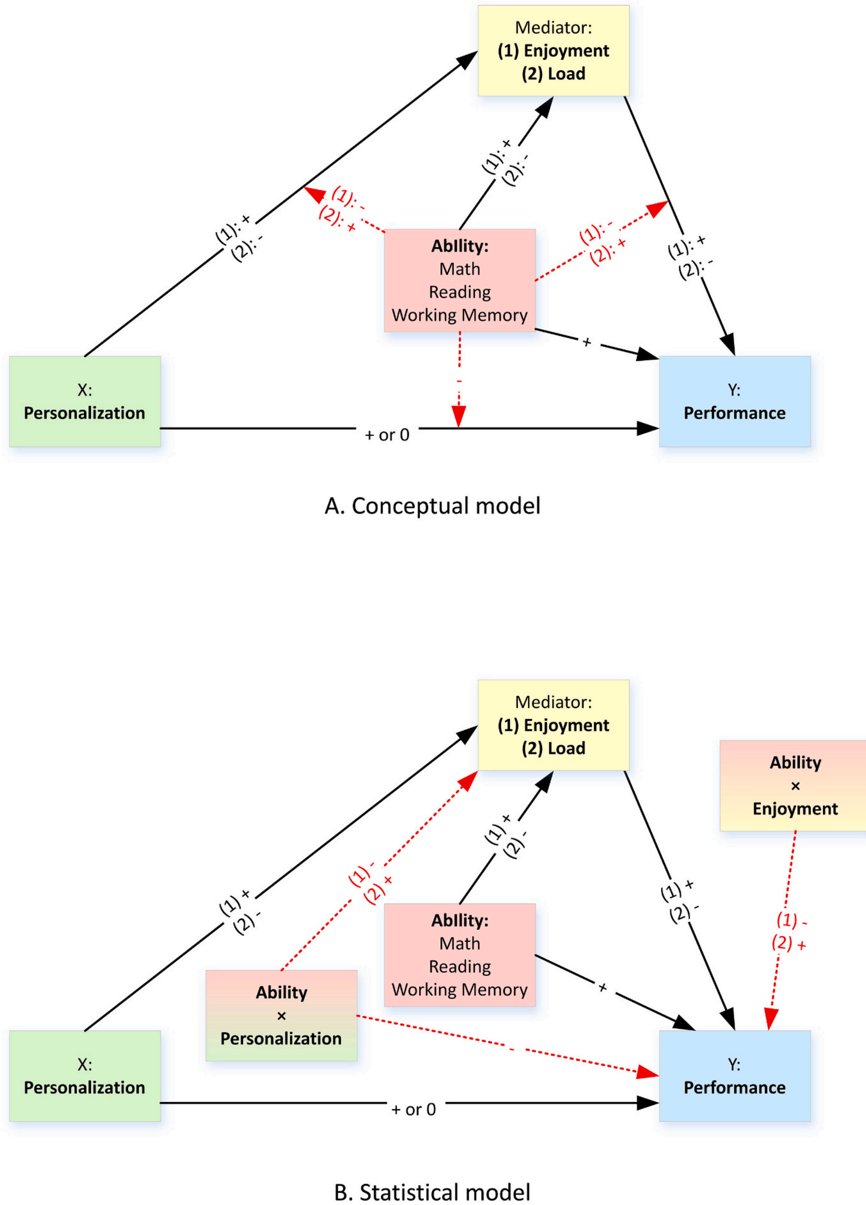

Fig. 1. Conceptual model (panel A) and the corresponding statistical model (panel B) of moderated mediation, with signs indicating the hypothesized directions of effects for (1) enjoyment and (2) cognitive load as mediator. Red, dotted arrows represent moderation effects. Path names are identical in both models. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

personalization on performance, which may be so strong that there is no main effect of personalization. Note that expected signs for enjoyment and cognitive load differ because we expect enjoyment to be positively related to personalization and performance, while we expect load to be negatively related. We explain the expected pattern for each proposed mediation mechanism.

1.4.1. Motivation as mediator

We expect personalization to increase enjoyment (path A), which leads to better performance (path B). Increased enjoyment may thus explain a possible direct positive effect of personalization on performance (path C). We also expect ability to predict enjoyment (path D) (Frenzel et al., 2007; Murayama et al., 2013; Prast et al., 2018) and performance (path E). In the moderation part of the model, we expect personalization to be especially effective for lower-ability students (path C_{mod}). This may be explained because students of lower ability will show a stronger increase in enjoyment: a negative moderation effect of ability on the path from personalization to motivation (paths A_{mod}). We also expect the effect of enjoyment on performance to be stronger in students of lower ability (path B_{mod}), as they may need a higher level of enjoyment to put in effort into solving the math problems. We expect this pattern for different abilities necessary for mathematical word problem solving: math, reading and WM.

1.4.2. Cognitive load as mediator

For cognitive load, the argument is similar, although the signs are reversed. We expect personalization to reduce cognitive load (path A), while cognitive load is a negative predictor of performance (path B), which may explain a positive effect of personalization on performance (path C). Ability will be predictive of word problem performance (path E), (partly) because students' lower abilities make them experience a higher cognitive load (path D). Again we expect moderation: stronger relations in lower-ability students. First, as in the previous model, we expect that personalization will increase performance more in lower-ability students (path C_{mod}), partially because cognitive load reduces more in these students (path A_{mod}). This is reflected in a positive moderation effect: personalization may free up the limited resources of students with lower ability more strongly. Finally, we expect the effect of reduced cognitive load on performance to be stronger for lower-ability students (path B_{mod}): as these students are more likely to suffer cognitive overload, the reduction caused by personalization may be especially relevant for them.

1.5. The present studies

We report on two studies that examined the hypotheses that (a) personalization effects are mediated through increased enjoyment and reduced cognitive load, and (b) these mediation effects are moderated by math, reading and WM ability. Although motivation and cognitive load have been put forward to explain general and differential effects of personalization (Harter & Ku, 2008; Ku & Sullivan, 2000), we have no knowledge of any studies examining such a conditional process model systematically. In the present studies, we focus on 6th grade students, since research shows that motivation for mathematics declines towards the end of primary schools (Gottfried, 2001; Gottfried et al., 2009) making this an important period for motivational interventions (Cleary & Chen, 2009; Gottfried et al., 2009). Also, selecting one grade makes it more feasible to select appropriate math problems. The studies were approved by the ethics committee of the Social and Behavioral Sciences Faculty, Utrecht University.

2. Study 1

2.1. Method

2.1.1. Design

A between-subjects experimental design with two conditions (personalized versus control) was used. In the personalized condition, children solved problems about buying items that they had previously selected in an interest inventory. Furthermore, the child's name was inserted. In the control condition, children solved numerically identical problems containing standard items deemed not interesting for children, and a standard name was used.

2.1.2. Participants

Students from eleven 6th grade classes of nine primary schools in The Netherlands participated. Parents of 241 children received written information. A passive consent procedure was used and the parents of three children declined consent, leaving a total of 238 participants (54.6% boys, mean age: 11.9 years). Participants were matched on gender, WM, math, and reading ability and then the pair was randomly divided over the two conditions. However, a few children had data missing on the interest inventory and these were placed in the control condition. Children in the control condition ($n = 125$) and the personalized condition ($n = 113$) did not differ with regard to gender, $\chi^2(1) = 0.019$, $p = .890$ or measures of WM, $t(203) = 0.504$, $p = .615$ and $t(207) = 0.521$, $p = .303$, math, $t(215) = 0.354$, $p = .304$, and reading ability, $t(217) = -0.195$, $p = .924$ (see Table 1 for descriptives).

2.1.3. Measurements

2.1.3.1. Personalization experiment

2.1.3.1.1. Interest inventory. Children's interests were assessed with a computerized inventory with four lists of items in different prize categories (≤€10; €10–€20; €20–€50; ≥€50), each containing sixteen pictures (e.g., painting set, computer game). After inspecting toy catalogues, and consulting teachers and children, items were selected from different themes (e.g., sports, music, toys, clothing) to cover a wide variety of interests. In each list, children were asked to select three

Table 1
Descriptive statistics for Study 1.

	Control			Personalized		
	N	M	*SD*	*N*	M	*SD*
Visual-spatial WM[a]	102	0.80	0.12	103	0.79	0.10
Verbal WM[a]	105	0.61	0.12	104	0.60	0.13
Math	112	109.38	12.27	105	107.61	12.98
Reading	114	45.95	15.92	105	46.14	14.06
Word problem performance[a]	117	0.62	0.31	106	0.63	0.29
Enjoyment	117	3.46	0.94	106	3.57	0.79
Cognitive load	117	3.38	1.67	106	3.26	1.49
Interest	116	3.27	1.10	105	3.70	1.01

WM = working memory.
[a] Proportion correct.

items, resulting in twelve items of interest per participant. Appendix A shows the complete inventory.

2.1.3.1.2. Word Problem Performance. Three to four weeks after the interest inventory, students received a booklet containing eight mathematical word problems with percentages, one per page. Percentages was chosen as this domain was relatively recently introduced and variation was expected to be large. In the control condition, word problems involved buying items deemed uninteresting for children (see Appendix A for a full list of the word problems). In the personalized condition, the mathematical problems were identical but names and items were personalized based on the student's interest inventory. An example of a control (personalized) word problem is: "*Ethel (student's name) buys a ladder (computer game). The ladder (computer game) costs €34. Ethel (student's name) gets a 25% discount. How much does Ethel (student's name) have to pay?*". On each page, a section was presented as scratch paper. The proportion of correct answers was used as a final score.

2.1.3.1.3. Enjoyment. Enjoyment was assessed immediately after each word problem. Children rated "I found working on this math problem…." on a five-point Likert-scale ranging from 'not enjoyable at all' (1) to 'highly enjoyable' (5), accompanied by five icons with faces ranging from unhappy to happy, adopted from (Ainley et al., 2002). The statement was modified from a scale used to assess ongoing motivational appraisals (Niemivirta & Tapola, 2007). A mean enjoyment score over the 8 items was calculated. Internal consistency in the current sample was high (Cronbach's $\alpha = 0.92$).

2.1.3.1.4. Cognitive Load. Immediately after each word problem, students also answered a question adapted from Paas (1992): "How much effort did it cost you to solve this problem?" on a nine-point scale ranging from (1) very, very little effort, to (9) very, very high effort. Research has shown that cognitive load self-report ratings give a reliable and sensitive representation of objective cognitive load, as self-reported cognitive load is related to task complexity and novelty (Paas et al., 1994). A mean cognitive load score over the 8 items was calculated. Internal consistency in the current sample was high (Cronbach's $\alpha = 0.91$).

2.1.3.1.5. Interest. After completing all word problems, children answered the question: "To what extent did the problems in this booklet address things you like?" on a five-point scale ranging from (1) not at all to (5) very much. This served as a manipulation check of the personalization.

2.1.3.2. Ability measures

2.1.3.2.1. Working Memory. Two online computerized WM tasks suitable for self-reliant administration in the classroom were administered: the Lion game and the Monkey game. The Lion game is a visual-spatial complex span task, in which children recall the locations of colored lions (van de Weijer-Bergsma et al., 2015). In every item, eight lions of five different colors are consecutively presented in a 4 × 4 matrix at different locations for 2000 ms each. Children are asked to remember the last location where a lion of a certain color (e.g., red) appeared. The task consists of five levels of each four items, in which the number of colors – and hence, the number of locations - children have to remember and update increases from one to five. The Lion game has excellent internal consistency (Cronbach's α between 0.86 and 0.90), satisfactory test-retest reliability ($\rho = .71$) and good concurrent and predictive validity (van de Weijer-Bergsma et al., 2015).

The Monkey game is a backwards word recall task, in which children are presented with audio-recorded one-syllable words, and asked to recall the words in backwards order (van de Weijer-Bergsma et al., 2016). Children respond by clicking on the corresponding written words presented in a 3 × 3 matrix in backward order. The task consists of five levels of each four items, in which the number of words to be recalled increases from two words in level 1 to six words in level 5. The Monkey game has excellent internal consistency (Cronbach's α between 0.78 and 0.89) and shows good concurrent and predictive validity (van de Weijer-Bergsma et al., 2016).

In both games, no cut-off rules were applied. We scored the proportion of stimuli (lions or words) recalled correctly for each game. Then an aggregated WM score was constructed by first standardizing both scores and taking the mean of these standardized scores.

2.1.3.2.2. Reading. The criterion-based Cito Reading Comprehension Test (Feenstra et al., 2010; Weekers et al., 2011) scores from the previous year were obtained from the schools. The test consists of 55 multiple choice questions about different reading passages. Raw scores are converted into IRT-based ability scores. Validity and reliability are satisfactory (Cronbach's α ranges from 0.84 to 0.93; Feenstra et al., 2010).

2.1.3.2.3. Mathematics. The criterion-based Cito Mathematics Test (Janssen et al., 2010) scores from the previous year were obtained from the schools. In the test, several math domains are covered, including computation, measuring length, time, and money. Raw test scores are converted into IRT-based ability scores (Janssen et al., 2010).

2.1.4. Procedure. Data were collected in two cohorts, one in spring 2017 and one in spring 2018, and then combined. In both cohorts, data were collected during two measurement occasions in a classroom setting. At the first measurement occasion, the interest inventory and two WM tests were administered. Also, math and reading test scores from the previous school year (i.e., grade 5) were collected from teachers. At the second measurement occasion, the personalization experiment was conducted with the booklet containing mathematical word problems as well as the enjoyment, cognitive load and interest questions.

2.1.5. Data analysis. First it was tested with an independent-samples *t*-test whether the word problem booklet matched the interests of students more in the personalized condition compared to the control condition. Cohen's *d* is reported as a measure of effect size, where 0.2, 0.5 and 0.8 correspond to small, medium and large effect sizes respectively (Cohen, 1988).

Then, to test moderated mediation effects, we estimated regression-based conditional process models using the PROCESS macro for SPSS, version 3.3 (Hayes, 2018), as depicted in Fig. 1B, in which all three paths of the mediation model were moderated by ability (Model 59; Hayes, 2018). The heteroscedasticity-consistent inference estimator HC3 (Davidson-MacKinnon) was used, as recommended by Hayes and Cai (2007).

We had three different moderators: math ability, reading ability, and WM ability, and two mediators: enjoyment and cognitive load. Because incorporating all these in a single model would result in an overly complex model, each combination of moderator and mediator was run once, which yielded six conditional process models. All continuous predictors were standardized before the analysis. The coefficients thus represent standardized effects. Main effects of Condition reflect the standardized difference between the two conditions, and interaction effects with Condition reflect the difference in standardized effects between the conditions. We report the coefficients (*B*) and their standard error (SE).

Throughout the study, the alpha level was set at .05. The indirect effect (the effect of the intervention on performance through the respective mediators enjoyment and load), based on the 95% confidence interval (CI) derived from 5000 bootstrap resamples, is significant at the 5% level if the CI does not include zero. Moderation effects with a *p*-value $\leq .05$ were further probed with the Johnson-Neyman technique to reveal at which levels of the moderator there was a significant effect (Hayes, 2018).

2.1.5.1. Missing values. Of the 238 children, 14 children were absent at the experiment (due to illness or dentist visits) and 1 child did not participate because of severe mathematical problems. These children were excluded from the analyses. Of the remaining 223 children (personalized: $n = 106$; control: $n = 117$), 11 children had missing data on the interest inventory and both WM assessments due to illness. Due to technical problems data were missing for 6 more children on the Lion game, and for 3 children on the Monkey game; for these children the remaining WM task score was used as final score. Data from mathematical and reading ability were missing for 6 and 4 children, respectively. Two children had not answered the interest question. Each analysis was performed with the children with data for the variables included. Some children had partial missing data: 4 children had not answered a math problem, which was scored as incorrect; 8 children had not answered both the enjoyment and cognitive load questions for one ($n = 7$) or two ($n = 1$) items, and two had not answered only a enjoyment ($n = 1$) or cognitive load ($n = 1$) question; mean scores were calculated with the remaining items.

2.2. Results

Descriptive results are presented in Table 1.

2.2.1. Interest

An independent samples *t*-test showed that children in the personalized condition rated the word problems as addressing their interest more than children in the control condition, $t(219) = -3.07$, $p = .002$, Cohen's $d = 0.41$, a small to medium effect. This confirms that the personalization procedure had worked as intended.

2.2.2. Conditional process models

Fig. 2 presents the results from the six conditional process models, with separate figures for motivation (a) and cognitive load (b) as mediators. For each mediator, three different models were run with the three respective moderators (math ability, reading ability, WM ability). The resulting coefficients are presented together in each figure.

The results show that personalization affected neither performance nor both mediators: enjoyment or cognitive load. Ability moderated the effect of personalization on performance only in one of the six models: when mathematics ability was the moderator and cognitive load the mediator, the effect of personalization was stronger in students of lower ability.

The rest of the analysis shows more significant relations. Ability significantly predicted performance: math ability was a medium to strong predictor, and reading and WM ability were both weak to medium-sized predictors. Both mediators were also significant: enjoyment positively predicted performance and cognitive load negatively predicted performance. The effects of both were similar in size: weak to medium in the models with math ability, and medium to strong in the models with reading and WM ability. Moreover, math, reading and WM ability all significantly moderated the latter two relations, such that the effects of the mediators were stronger in students of lower ability, with one exception: reading ability did not moderate the effect of cognitive load on performance. Probing with the Johnson-Neyman technique

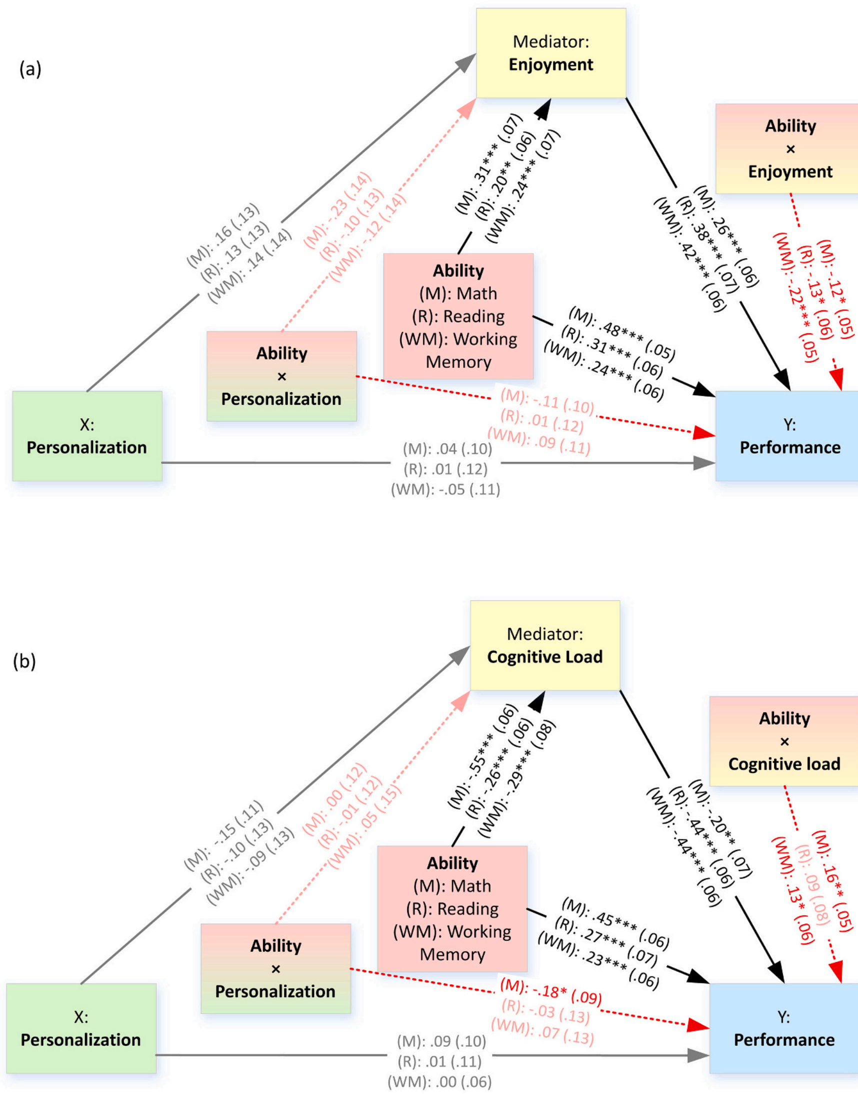

Fig. 2. Results from the six conditional process models of Study 1, with mediators Enjoyment (a) and Cognitive load (b). Each figure contains the results from three different analyses, each with a different moderator (math, reading, and WM ability). Red, dotted arrows represent moderation effects.
$^*p < .05$, $^{**}p < .01$, $^{***}p < .001$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

showed that the effect of enjoyment on performance was significant for students with a (standardized) math ability of $+0.88$ or lower (82.0% of the sample), a reading ability of $+1.28$ or lower (88.1%) and a WM ability of $+1.14$ or lower (89.6%). The effect of cognitive load on performance was significant for students with a math ability of -1.36 or lower (8.3%) and a WM ability of -1.55 or lower (1.56%).

2.3. Discussion study 1

The results of Study 1 show that the problems in the personalized condition matched the children's interest more than in the control condition. Nevertheless, effects of personalization on math performance were only weak, with no main effect, and a moderated effect of ability on performance in only one instance: when math ability was the moderator in the model with cognitive load as mediator. As expected, personalization was then more effective in students of low ability.

Despite this underwhelming overall effect, however, the two proposed mechanisms that we identified, namely enjoyment and cognitive load, did explain differences in math performance themselves, also when correcting for ability. Furthermore, these relations between enjoyment and cognitive load on the one hand and performance on the word problems on the other hand (path B in a mediation model, see also Fig. 1) were moderated by math, reading and WM ability. This pattern of results suggests that we identified promising mediating and moderating candidates, but that the personalization we applied in our study was not sufficiently successful to enhance enjoyment or reduce cognitive load. There are two possible explanations for this. First, Walkington and Bernacki (2018) indicated that a certain level of difficulty is desirable for performance, and that personalization may be redundant and even increases cognitive load when problems are too easy. The accuracy broken down by math problem ranged from 48.4% (most difficult problem) to 79.8% (easiest problem), indicating that they may have been too easy to render some of the hypothesized effects. Another possible explanation is related to the way we personalized the problems. Possibly the depth, grain size and/or ownership in the personalization was not strong enough. In the present study, personalization was rather superficial, coarse-grained and students may have experienced limited ownership. Although increasing the depth of personalization may not be easily feasible in daily educational practice, grain size and ownership may be increased more easily. With this aim, we set up a second study.

3. Study 2

Similar to Study 1, in Study 2 we examined whether the effects of personalization are mediated by enjoyment and cognitive load, and whether these effects are moderated by individual differences in mathematical ability, reading ability and WM ability.

While the study design and measures are identical to Study 1, there are two important differences. We aimed at a smaller grain-size and greater ownership in the personalized problems: students were provided with a larger variety of pre-defined choices and open fields so students could add their own choices. Furthermore, an overarching theme with a narrative was chosen; celebrating your birthday with friends, to create a higher level of interest, including a social element, which is highly important for adolescents (Forbes & Dahl, 2010; Nelson et al., 2016). Second, to take problem difficulty into account, word problems with more varying difficulty levels were selected.

3.1. Method

3.1.1. Design

A between-subjects experimental design, identical to Study 1, with two conditions (personalized versus control) was used.

3.1.2. Participants

Children from six 6th grade classes from five primary schools in The Netherlands participated. Parents of 151 children received written information. A passive consent procedure was used and the parents of two children declined participation of their child, leaving 149 participants (43.6% boys, mean age: 12.0 years). Participants were matched on gender, WM, math, and reading ability; the pair was randomly divided over the two conditions. Children in the control condition ($n = 77$) and the personalized condition ($n = 72$) did not differ with regard to gender, $\chi^2(1) = 0.003$, $p = .960$, or measures of WM, $t(133) = 0.660$, $p = .242$ and $t(133) = -0.609$, $p = .480$, math $t(135) = -0.107$, $p = .604$, and reading ability, $t(136) = -0.380$, $p = .917$ (see Table 2 for descriptives).

3.1.3. Measurements

3.1.3.1. Personalization experiment

3.1.3.1.1. Interest Inventory. Children's interests were assessed with a computerized inventory consisting of different questions on how they would want to celebrate their birthday. Questions were aimed at desired presents, friends they wanted to invite, their preferred activity, and their preferred food and beverage. Most items consisted of both a predefined list with items from different themes (e.g., sports, music, toys), and an open answer format to include additional interests. See Appendix B for the full inventory.

3.1.3.1.2. Word Problem Performance. Three to four weeks after the interest inventory, children received a booklet containing 10 mathematical word problems (5 easy and 5 difficult problems), one per page, tapping into different domains and operations, such as addition, multiplication, percentages. All problems were created around the theme of celebrating a birthday. In the control condition, items were selected that we deemed to lay outside the interest areas of the children (see Appendix B for a list of word problems in the control condition). In the personalized condition, word problems were identical but names, items and places came were personalized according to each student's interest inventory. On each page, a section was presented as scratch paper. The proportion of correct answers was used as a final score.

Table 2
Descriptive statistics for Study 2.

	Control			Personalized		
	N	M	*SD*	*N*	M	*SD*
Visual-spatial WM	66	.83	.09	69	.81	.12
Verbal WM	67	.64	.11	68	.65	.12
Math	69	110.81	11.83	68	111.01	10.37
Reading	70	50.64	15.38	68	51.66	16.13
Word problem performance[a]	72	.52	.21	69	.53	.21
Enjoyment	72	3.07	.79	69	3.26	.80
Cognitive load	72	3.80	1.29	69	3.68	1.32
Interest	71	3.09	.91	68	3.74	.82

WM = working memory.
[a] Proportion correct.

3.1.3.1.3. Enjoyment. *Enjoyment* and *Cognitive Load* were repeatedly assessed immediately after each word problem, using the same questions as used in Study 1. Also identical to study 1, *interest* was assessed with a single question after completing all word problems, as a manipulation check of the personalization.

3.1.3.2. Ability measures.

The same measures that had been used in Study 1 were used to assess *working memory* (i.e., the Lion game and Monkey game), *reading ability* (i.e., Cito Reading Comprehension Test) and *math ability* (i.e., Cito Mathematics test).

3.1.3.3. Design & procedure.

Data were collected at two measurement occasions during spring 2018 in a classroom setting. Identical to study 1, the first measurement occasion included the interest inventory and two WM tests and the collection of math and reading test scores from the previous school year (i.e., grade 5). During the second measurement occasion, the personalization experiment included the booklet with mathematical word problems and questions regarding enjoyment, cognitive load and interest.

3.1.4. Data analysis

Data analysis was identical to study 1: a manipulation check, then moderated mediation models.

3.1.4.1. Missing values.

Of the 149 children, 8 children were absent at the post-test experiment (due to illness or holiday leave). Data for these children were removed from the dataset. Of the remaining 141 children (personalized: $n = 69$; control: $n = 72$), 4 children had been absent on the pretest due to illness and had missing data on both WM assessments. Due to technical problems data were missing for 2 children on the Lion game, and for 2 children on the Monkey game; for these children the score of the other WM task was used as final score. Data on mathematical and reading ability tests were missing for 4 and 3 children, respectively. Two children had not answered the interest question. Each analysis was performed with children who had complete data for the variables included. Missing data on one ($n = 15$), two ($n = 5$) or three ($n = 4$) of the ten math problems, were scored as incorrect. Two children had not answered both the enjoyment and cognitive load questions for one ($n = 1$) or two ($n = 1$) math problems, and mean scores were calculated with the remaining items.

3.2. Results

Descriptive results are presented in Table 2.

3.2.1. Manipulation check

An independent samples *t*-test showed that children in the personalized condition rated the word problems as addressing their interest more than children in the control condition, $t(137) = -4.40$, $p < .001$, Cohen's $d = 0.75$, a medium to large effect. The personalization manipulation had thus worked as intended.

3.2.2. Conditional process models

Fig. 3 presents the results from the six conditional process models, with separate figures for enjoyment (a) and cognitive load (b) as mediators. For each mediator, three different models were run with the three respective moderators (math ability, reading ability, WM ability), The resulting coefficients are presented together in each figure.

The results show that personalization did not significantly affect performance, motivation, or cognitive load. Student ability did predict

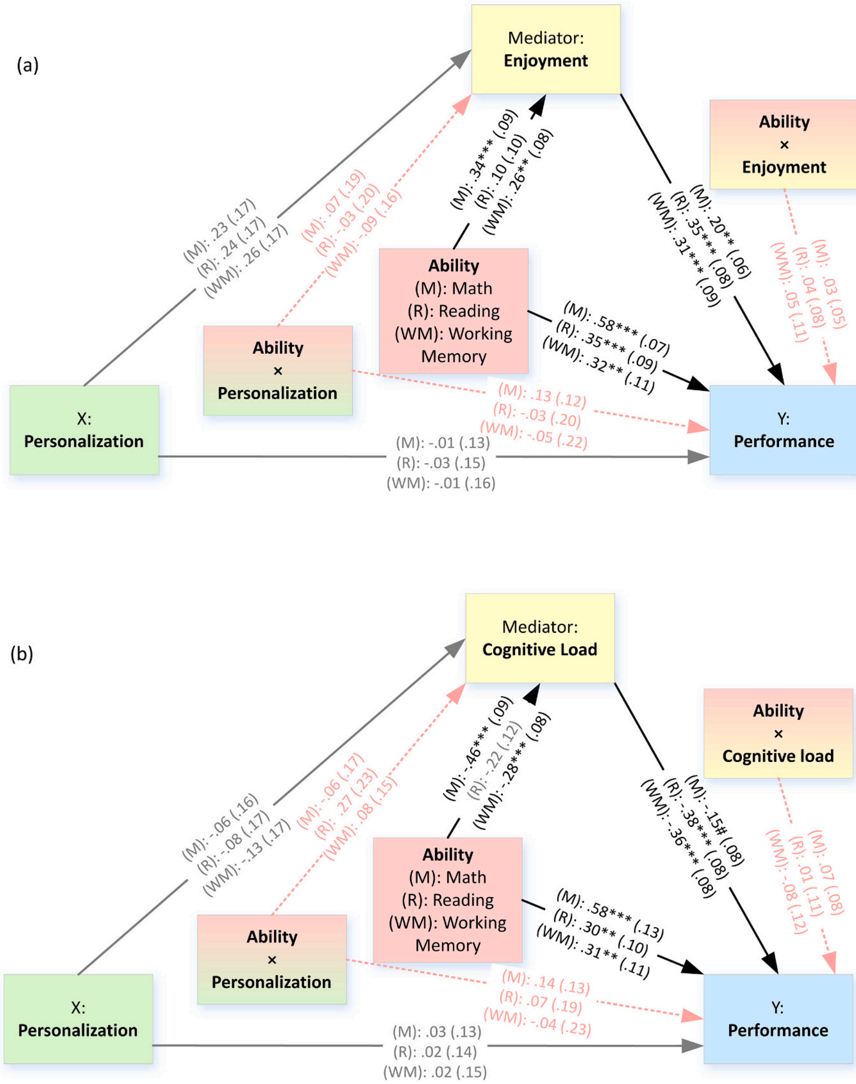

Fig. 3. Results from the six conditional process models of Study 2, with mediators Enjoyment (a) and Cognitive load (b). Each figure contains the results from three different analyses, each with a different moderator (math, reading, and WM ability). Red, dotted arrows represent moderation effects.
$^*p < .05$, $^{**}p < .01$, $^{***}p < .001$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

both mediators: math and WM (but not reading) ability significantly predicted higher enjoyment and lower cognitive load, with a medium effect size. Ability also significantly predicted word problem performance: the effect of math ability was strong, the other two abilities (WM and reading) of medium size. Both mediators also significantly predicted word problem performance: the effect was small to even non-significant in the models with math ability, and medium-sized in the models with reading and WM ability. Ability did not moderate the predictive effects of enjoyment and cognitive load on word problem performance.

To test the effect of problem difficulty, we also ran the conditional process models with the easy and the difficult math problems separately, but these analyses did not reveal a different pattern of results. Of all 36 possible moderation effects (2 difficulty levels $\times$ 2 mediators $\times$ 3 moderators $\times$ 3 moderated paths), only 2 were significant at the 0.05 level, one of which was in the expected direction and the other in the reverse direction. We take these results as reflecting chance, and do not report them in detail here.

3.3. Discussion study 2

The results of Study 2 show that the problems in the personalized condition matched the children's interest more than in the control condition. However, further beneficial effects of personalization on enjoyment, cognitive load or performance were not found, not even for participants with lower math, reading or WM ability. In line with Study 1, enjoyment and cognitive load significantly predicted math performance, even when correcting for ability. However, in contrast to study 1, these relations were not moderated by ability.

4. General discussion

In two studies with different degrees of personalization, we examined whether context personalization of mathematical word problems enhanced performance by increasing enjoyment and reducing cognitive load in 6th grade students. Furthermore, we examined whether math, reading and WM ability moderated these relations. We expected especially lower ability students to benefit from personalization, while effects may even be counterproductive in higher ability students.

4.1. Effects of personalization

In both study 1 and 2, we found no effect of our interventions on performance. Since context personalization may not be effective for all students, we did not necessarily expect a main effect. However, contrary to our expectations, we did not find that ability moderated the effect of personalization either. That is, only in study 1, and only in the model with cognitive load as a mediator, math ability weakly moderated the effect of personalization: students with a very low math ability benefited from personalization, in line with previous studies (Harter & Ku, 2008; Ku & Sullivan, 2000; Walkington et al., 2013). Because we only found a moderation effect of math ability in one out of four analyses, we believe this is likely to be a chance finding. Moreover, the other two abilities, reading and WM, never significantly moderated this relationship. Taken together, these findings show that the personalization manipulation did not specifically benefit children with math, reading or working memory difficulties. Moreover, contrary to our expectations, our interventions did not affect enjoyment or cognitive load. Although cognitive load has not been examined before, these findings are in contrast to findings from earlier studies, that showed increased motivation for personalized word problems (Akinsola & Awofala, 2009; Høgheim & Reber, 2015, 2017; Ku & Sullivan, 2000, 2002; López & Sullivan, 1992; Walkington & Bernacki, 2014). There are several possible explanations for the failure to find an effect of context personalization on performance, enjoyment and cognitive load in our studies.

4.1.1. Principles of personalization

One explanation relates to the principles of depth, grain size and ownership. First, the depth of the personalization was probably too shallow. In both studies many problems contained total prices of items. While an item may match the students' interest, the problems may not have related to how they spontaneously use numbers related to their interests (Walkington & Bernacki, 2018). Such a match may connect more strongly to the students' interest, and thus enhance their motivation. However, this requires students to use mathematics spontaneously, which students with low math ability may avoid completely. Furthermore, it requires a far more profound preparation in creating the materials, as not only the items in the problem but the entire problems need to be personalized. Moreover, as the mathematics textbooks of the last decennia have changed towards the use of math problems that are grounded in familiar contexts, often deemed attractive for children, children may already be used to what could be considered group-level personalization. In a regular school context, with one teacher teaching many children, more in-depth personalization would be overly time consuming. Smart ICT solutions might be helpful but even this is complicated, as somehow the software needs to know how a child uses math outside school.

Second, the grain size of the personalization in our studies may have been too coarse, despite the addition of an open-ended answer format in the interest inventories. It is important to emphasize, however, that our manipulation did not increase cognitive load or decrease performance either. Walkington and Bernacki (2018) pointed out that shallow or coarse personalization may increase errors when details are irrelevant or redundant, and may thus add cognitive load and hamper performance. We did not find such a 'seductive detail' effect.

Third, the degree of ownership experienced by students may not have been high enough, as the children had no control over task selection. In Study 2 we tried to increase ownership by giving more (open) personalization choices. However, the task was still fixed, and the 3–4 weeks delay between the interest inventory and the math booklet may have reduced the experience of ownership. Walkington and Bernacki (2018) indicate that ownership is probably most effective when students play an active role, by being co-creators and incorporating their interest in the tasks themselves.

4.1.2. Problem solving steps

Our finding that neither performance nor cognitive load benefited from personalization may also be related to the fact that solving a problem consists of three steps: building a problem representation, selecting the calculations, and executing the calculations. Possible effects of personalization may, in theory, be found in all steps, especially in complex, abstract problems. Personalization can make a complex problem representation easier to understand and select the appropriate calculation procedure, and when this procedure is complex and novel, it may help students refer back to the context during problem execution to keep track of what they are doing. However, when calculations are easy and familiar, this may all be unnecessary. Because even the non-personalized contexts were likely familiar schemas for all children and the mathematical operations we apparently too easy, the required load to construct the problem from the context may have been too low to be further reduced. In the third step of executing the calculation, cognitive load may be higher for problems requiring difficult computational steps. However, personalization may not alleviate this type of load. Age could be a factor too: at a younger age, i.e., grades 1–3, performance differences between problems with and without a context are larger (Hickendorff, 2013a), perhaps because contextual information puts a higher load on WM at this age, and children are less used to word problems at this age. A similar study in younger children can shed more light on this issue. A limitation of the current studies is that we included grade 6 pupils only, which impacts the generalizability of the findings.

4.1.3. Practicing existing skills versus learning new skills

Another possible explanation for the lack of effects on performance or cognitive load may have to do with the difference between practicing existing skills and learning new ones. We only personalized math problems, not the instruction. This strategy allowed us to examine the direct relationship between cognitive load and motivation for the specific math problems and how these affected performance on these same problems. However, personalization may only be beneficial to math achievement or cognitive load when a new mathematical topic is introduced, requiring new conceptual or computational development. This is likely to be accompanied by a high level of complexity. Students may thus benefit much more from personalization during instruction than during practice. Personalization can then serve to reduce cognitive load while students are grounding new concepts. Indeed, several studies that found beneficial effects of personalization (Clinton & Walkington, 2019; Ku & Sullivan, 2002) applied personalization to the instructional phase as well.

4.1.4. Measurement of enjoyment and cognitive load

Our measure of enjoyment and cognitive load, both the contents and the timing, may also explain our findings. Motivation was operationalized as enjoyment and repeatedly assessed after each word problem, using the formulation "I found working on this math problem…" (not enjoyable at all – highly enjoyable). Possibly, asking this question in these specific words after each problem triggered participants to especially reflect on how much effort it took and how much they enjoyed executing the calculation ('…working on the math problem'), which was the same in both conditions, even though participants in the personalized condition may have found constructing the problem representation easier and more enjoyable. The results from the manipulation check, which was conducted after all problems had been solved, indicated that the personalized booklet did indeed match students' interest more, which is in line with other studies that found situational interest to increase (Bernacki & Walkington, 2014; Høgheim & Reber, 2015, 2017). The word 'enjoyable' may also have been too strong: the problems may have increased interest but that may not have made them enjoyable. Moreover, our operationalization of enjoyment as reflecting intrinsic motivation may have been too limited. We recommend the inclusion of a more comprehensive scale including several questions to assess intrinsic motivation after all problems have been solved for future studies. Furthermore, a clear definition of motivation and which aspects of motivation may be affected by personalization is recommended, as well as sharp attention for the formulation and timing of the self-report questions. Asking students different types of motivation questions may reflect broader aspects of motivation, making it a more sensitive measure. Several types of motivation can be distinguished (e.g., autonomous motivation, task motivation), as these types of motivation affect the strength of expected situational effects (Ryan & Deci, 2000).

Finally, we used a self-report measure of cognitive load, meaning that it is a subjective measure that may also be affected by a student's confidence in the correctness of their answer. A more objective measure, such as a physiological measure during the task (e.g., pupil dilation), may be preferable in future research.

4.2. Effects of enjoyment and load on performance

Although not a main objective, both studies revealed that enjoyment and cognitive load significantly predicted performance, when controlling for math, reading and WM ability respectively. In study 1, these effects were stronger for students with lower ability. (Jõgi et al., 2015; Logan et al., 2011). Our results also showed that lower-ability students, especially in math, experienced lower enjoyment and more cognitive load than students with a higher ability. We took these results to suggest that enjoyment and experienced cognitive load are especially important for students with lower ability (Jõgi et al., 2015; Logan et al., 2011). Furthermore, cognitive load and enjoyment may have reciprocal associations: because enjoyment makes it easier to put an effort into a task, students may experience a lower cognitive load. Conversely, when students experience cognitive overload, their motivation for the task may be greatly reduced. However, as discussed before, our personalization intervention was not strong enough and/or did not target the right aspects to increase enjoyment and decrease cognitive load sufficiently.

4.3. Implications

Our findings suggest that personalization may only be effective under specific circumstances, with task conditions and learner characteristics interacting. Moreover, the underlying mechanisms may differ in varying circumstances. For example, personalization may reduce cognitive load when complexity is high, such as during instruction of new mathematical concepts. Furthermore, for students with math difficulties, who need elaborated practice, personalization may increase their motivation for such repeated practice, but effects on achievement may not become immediately visible. Moreover, as the effectiveness of personalization may fade when it is repeatedly employed, personalization should be part of a larger motivation-enhancing repertoire. Future studies should 1) focus on examining motivation and cognitive load as explanatory mechanisms of context personalization during instructional activities introducing new mathematical concepts, 2) further examine interactions between learner characteristics and personalization dimensions (depth, grain size, ownership), and 3) try to replicate our finding that the benefits from higher enjoyment and lower experienced cognitive load are larger in low-ability students. This finding should be confirmed with an experimental design that is more effective in altering enjoyment and cognitive load than our personalization intervention was.

4.4. Conclusion

In the present study, personalization of math word problems did not benefit student performance, regardless of their math, reading or WM ability, nor did it affect enjoyment or cognitive load. Future studies should focus on further elucidating the when, for whom and why personalization affects mathematical learning, but also on ways to increase enjoyment and reduce cognitive load, especially for lower-ability students.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

None.

Acknowledgements

We thank all the teachers and students who gave their time and effort to participate in this study. The assistance of Sophie Westerink, Lieke Boon, Mirja Pleijsant, Marit van Hoek, Els van Dorp, Marlinde van Pijkeren, Wytske Huitema, Irene Sloeserwij, Lisa van Blokland, Maud Verstappen and Malou Roestenburg in the collection of data is gratefully acknowledged.

Appendix A. Interest items and mathematical word problems for Study 1

A.1. Interest items

	Category I ≤€ 10	Category II € 10 to € 20	Category III € 20 to € 50	Category IV ≥€ 50
Items in interest inventory (used in personalized condition)	Bracelet	Football	Skateboard	Bicycle
	Phone case	Parfum	Zoo ticket	Drone
	Nail polish	Movie subscription	Hockey stick	Coat
	Swimming pool ticket	Lava lamp	Lego set	Tablet
	Selfie stick	Book	Computer game	Skeelers
	Goggles	Movie DVD	Make up set	Sneakers
	Comic book	Cap	Board game	Guitar
	Puzzle book	Chess board	Painting set	Action camera
	Sunglasses	Shoulderbag	Clothes	Hoverboard
	Baking set	Shawl	Back pack	Game console
	Colored pencils	Music subscription	Microscope	Wireless speaker
	Beanie	Science set	Dartboard	Trampoline
	Cinema ticket	Puzzle	Theme park ticket	Mobile phone
	Body cream	Basketball ring	Children's encyclopedia	Stunt scooter
	Music buds	Lasergame ticket	Radio controlled car	Keyboard
	Phone credit	Karting ticket	Wireless headphones	BMX bicycle
Control items	Cans of soup	Kettle	Ladder	Dining table
	Bottle of vitamins	Pans	Vacuum cleaner	Washing machine

A.2. Mathematical word problems for control condition and difficulty level (% correct)

Word problems		% correct
1.	Ethel buys two cans of soup. Two cans of soup normally cost € 6,-. Ethel has to pay € 4,-. What percentage is the discount that Ethel gets?	48.4%
2.	Ethel buys a jar of vitamin pills. There are 75 jars of pills in the store. There are 15 jars on the bottom shelf. What percentage of the jars is on the bottom shelf?	71.7%
3.	Ethel buys a water cooker. The water cooker costs € 18,-. Ethel gets a 30% discount. How much does Ethel have to pay?	54.7%
4.	Ethel sees an offer for pans. One pan costs € 13,-. Ethel gets a 50% discount on the second pan. How much does Ethel have to pay when she buys 2 pans?	79.8%
5.	Ethel buys a ladder. The ladder costs € 34,-. Ethel gets a 25% discount. How much does Ethel have to pay?	49.3%
6.	Ethels buys a new vacuum cleaner. The vacuum cleaner costs € 40,-. Ethel gets a 60% discount. How much does Ethel have to pay?	62.8%
7.	Ethels buys a dining table which is normally priced € 180,-. Ethel gets a 15% discount. How much does Ethel have to pay?	63.2%
8.	Ethels buys a new washing machine. The washing machine costs € 260,-. Ethels gets a 20% discount. How much does Ethel have to pay?	70.9%

Appendix B. Interest items and mathematical word problems for Study 2

B.1. Interest items

	Candy	Presents category up to € 10	Presents category € 10 to € 50	Presents category € 50 to € 100	Birthday pie	Birthday activity	Four friends
Items in interest inventory (used in personalized condition)	Predefined and open choice	Predefined and open choice	Predefined and open choice	Predefined and open choice	Predefined and open choice	Predefined	Open choice
	Marshmallow	Emoji erasers	Xbox giftcard	Trampoline	Apple pie	Lasergame	
	Candy hearts	Gel pen set	Table tennis set	Virtual reality glasses	Strawberry pie	Efteling (theme park)	
	Gummy bears	Magician set	Sun glasses	Beanbag	Lemon pie	Cook workshop	
	Mints	Pocket knife	Telephone case	Stunt scooter	Cream cake	Movie theatre	
	M&M’s	Squeezies	Card game	Drum set	Cheesecake	Circus workshop	
	Easter eggs	Make-up	Lava lamp	Karaoke set	Cake with candied fruit	Escape room	
	Salty licorice	Brain puzzle	Perfume	Sneakers	Mocca pie	Museum	
	Cola bottles	Monster slime	Chess board	Smart watch	Chocolate pie	Trampoline park	
	Jelly frogs	Selfiestick	Footbal	Inflatable familypool	Red velvet pie	Dance workshop	
	Sour mats	Bracelet	Music earbuds	Hoverboard	Mixed fruit pie	Horse riding class	
	Chewing gum ball	Yo-yo	Microscope	Computergame	Yogurtpie	Hellendoorn (theme park)	
	Winegum	3d snake puzzle	Puzzle	Television		Arts & crafts workshop	
	Dew drops	Fart pillow	Playstation giftcard	Drone		Zoo	
	Sweet licorice	Fidgetspinner	Diving mask	Keyboard		Carting rink	
	Toffee	Frisbee	Badminton set	Lego set de Luxe			

(*continued on next page*)

(*continued*)

	Candy	Presents category up to € 10	Presents category € 10 to € 50	Presents category € 50 to € 100	Birthday pie	Birthday activity	Four friends
						Walibi (theme park)	
		Unicorn cup	UNO card game	Buildingset		Subtropical swimming pool	
		Ant hotel	Book	Action camera		Bowling alley	
		Rubik's cube	Necklace	Skeelers		Duinrell (theme park)	
		Nail polish	Plasma ball	3D pen		Climbing	
		Puzzle book	Comic book	Gituar			
			Cap	Painting set			
			Scarf	Coat			
			Socker table	Remote controlled car			
				Wireless speaker			
				Telescope			
				Hockey stick			
				Airhockey table			
Control items	Carrots & tomatoes	Socks	Broom & bucket	Vacuum cleaner	Pasta & vegetable pie	City museum	Low frequent names

B.2. Mathematical word problems for control condition and difficulty level (% correct)

Word problem		% correct
1.	It's Ethel's birthday and she will buy treats. Ethel chooses carrots and tomatoes. She weighs 394 grams of vegetables. It costs € 5,97. She estimates how much 600 grams will cost. How many euro's will that cost approximately? Give your answer in whole numbers.	78.0%
2.	For her treat, Ethel divies 72 carrots and 72 tomatoes into 24 bags. How many vegetables will be in each bag?	80.1%
3.	Ethel invites Alfred, Herbert, Bertha and Eunice to her birthday. They will visit the city musuem with the five of them. Entrance fee is € 18,- per person. She gets a 15% discount on the total amount. How much does Ethel have to pay?	45.4%
4.	Bertha goes shopping for a present for Ethel. Bertha wants to buy 6 pairs of socks en compares prices at 2 shops. At one shop, she can buy 2 pairs for € 6,-. At the other shop, she can buy 2 pairs for € 8,-. How much euro is the difference in price when buying 6 pairs?	83.7%
5.	Together, Herbert and Alfred buy four presents for Ethel. Herbert buys a broom for € 9,80 and a bucket for € 10,60. Alfred buys a pen for € 3,- and a flash light for € 6,80. They want to share the costs evenly. How much does Alfred have to pay Herbert	25.5%
6.	On her birthday, Ethel and her friends travel by cart o the city museum. The distance between the house and the museum is 28 kilometres. They drive at an average speed of 90 km per hour. How much time in minutes do they travel? Give your answer in whole minutes.	24.8%
7.	During the party, they eat pasta and vegetable pie. At the end of the party, there is 2/3 of the pie left. They share it evenly with the five of them. Which part does each get?	19.1%
8.	Ethel wants to have a vacuum cleaner. It cost € 75,-. Family and friends have given her € 30,- in total for her birthday. She also saves € 7,50 form her allowance each week. How many weeks does she need to save her allowance to be able to buy the vacuum cleaner?	75.9%
9.	When Ethel walks by the shop she notices there is a sale. The vaccum cleaner cost € 75,- orginally, but now it cost € 60,-. What percentage discount does she get?	68.1%
10.	Now that the vacuum cleaner cost € 60,-, Ethel checks her money-box to see if she has enough money to buy it. There are 120 coins in her money-box: 5% are coins of 2 euro 35% are coins of 1 euro 15% are coins of 50 cent 45% are coins of 20 cent How much money has Ethel left, after she buys the vacuum cleaner?	19.1%

References

Ainley, M., Hidi, S., & Berndorff, D. (2002). Interest, learning, and the psychological processes that mediate their relationship. *Journal of Educational Psychology, 94*(3), 545–561. https://doi.org/10.1037//0022-0663.94.3.545.

Akinsola, M. K., & Awofala, A. O. A. (2009). Effect of personalization of instruction on students' achievement and self-efficacy in mathematics word problems. *International Journal of Mathematical Education in Science and Technology, 40*(3), 389–404. https://doi.org/10.1080/00207390802643169.

Bates, E. T., & Wiest, L. R. (2004). Impact of personalization of mathematical word problems on student performance. *Mathematics Educator, 14*(2), 17–26. http://eric.ed.gov/?id=EJ848502.

Bernacki, M., & Walkington, C. (2014). The impact of a personalization intervention for mathematics on learning and non-cognitive factors. *CEUR Workshop Proceedings, 1183*, 80–87 (doi:urn:nbn:de:0074-1183-1).

Cakir, O., & Simsek, N. (2010). A comparative analysis of the effects of computer and paper-based personalization on student achievement. *Computers and Education, 55* (4), 1524–1531. https://doi.org/10.1016/j.compedu.2010.06.018.

Cleary, T. J., & Chen, P. P. (2009). Self-regulation, motivation, and math achievement in middle school: Variations across grade level and math context. *Journal of School Psychology, 47*(5), 291–314. https://doi.org/10.1016/j.jsp.2009.04.002.

Clinton, V., & Walkington, C. (2019). Interest-enhancing approaches to mathematics curriculum design: Illustrations and personalization. *Journal of Educational Research, 112*(4), 495–511. https://doi.org/10.1080/00220671.2019.1568958.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Cordova, D. I., & Lepper, M. R. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of*

Educational Psychology, 88(4), 715–730. https://doi.org/10.1037/0022-0663.88.4.715.

Davis-Dorsey, J., Ross, S. M., & Morrison, G. R. (1991). The role of rewording and context personalization in the solving of mathematical word problems. *Journal of Educational Psychology, 83*(1), 61–68. https://doi.org/10.1037/0022-0663.83.1.61.

Feenstra, H., Kamphuis, F., Kleintjes, F., & Krom, R. (2010). *Wetenschappelijke verantwoording Begrijpend lezen voor groep 3 tot en met 6 [Scientific justification Reading comprehension test for grade 1 to 4]*. Cito.

Forbes, E. E., & Dahl, R. E. (2010). Brain and cognition pubertal development and behavior : Hormonal activation of social and motivational tendencies. *Brain and Cognition, 72*(1), 66–72. https://doi.org/10.1016/j.bandc.2009.10.007.

Frenzel, A. C., Thrash, T. M., Pekrun, R., & Goetz, T. (2007). Achievement emotions in Germany and China: A cross-cultural validation of the academic emotions questionnaire-mathematics. *Journal of Cross-Cultural Psychology, 38*(3), 302–309. https://doi.org/10.1177/0022022107300276.

Friso-van den Bos, I., van der Ven, S. H. G., Kroesbergen, E. H., & van Luit, J. E. H. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review, 10*(0), 29–44. https://doi.org/10.1016/j.edurev.2013.05.003.

Gottfried, A. E. (2001). Continuity of academic intrinsic motivation from childhood through late adolescence: A longitudinal study. *Journal of Educational Psychology, 93*(1), 3. https://doi.org/10.1037/0022-0663.93.1.3.

Gottfried, A. E., Marcoulides, G. A., Gottfried, A. W., & Oliver, P. H. (2009). A latent curve model of parental motivational practices and developmental decline in math and science academic intrinsic motivation. *Journal of Educational Psychology, 101*(3), 729–739. https://doi.org/10.1037/a0015084.

Guida, A., Tardieu, H., & Nicolas, S. (2009). The personalisation method applied to a working memory task: Evidence of long-term working memory effects. *European Journal of Cognitive Psychology, 21*(6), 862–896. https://doi.org/10.1080/09541440802236369.

Guida, A., Tardieu, H., le Bohec, O., & Nicolas, S. (2013). Are schemas sufficient to interpret the personalization effect? Only if long-term working memory backs up. *Revue Europeenne de Psychologie Appliquee, 63*(2), 99–107. https://doi.org/10.1016/j.erap.2012.09.006.

Hanushek, E. A., & Woessmann, L. (2010). *The high cost of low educational performance: The long-run economic impact of improving PISA outcomes*. OECD.

Harter, C. A., & Ku, H. Y. (2008). The effects of spatial contiguity within computer-based instruction of group personalized two-step mathematics word problems. *Computers in Human Behavior, 24*(4), 1668–1685. https://doi.org/10.1016/j.chb.2007.06.006.

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.

Hayes, A. F., & Cai, L. (2007). Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation. *Behavior Research Methods, 39*(4), 709–722. https://doi.org/10.3758/BF03192961.

Hickendorff, M. (2013a). The effects of presenting multidigit mathematics problems in a realistic context on sixth graders' problem solving. *Cognition and Instruction, 31*(3), 314–344. https://doi.org/10.1080/07370008.2013.799167.

Hickendorff, M. (2013b). The language factor in elementary mathematics assessments: Computational skills and applied problem solving in a multidimensional IRT framework. *Applied Measurement in Education, 26*(4), 253–278. https://doi.org/10.1080/08957347.2013.824451.

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4.

Høgheim, S., & Reber, R. (2015). Supporting interest of middle school students in mathematics through context personalization and example choice. *Contemporary Educational Psychology, 42*(April), 17–25. https://doi.org/10.1016/j.cedpsych.2015.03.006.

Høgheim, S., & Reber, R. (2017). Eliciting mathematics interest: New directions for context personalization and example choice. *Journal of Experimental Education, 85*(4), 597–613. https://doi.org/10.1080/00220973.2016.1268085.

Janssen, J., Verhelst, N., Engelen, R., & Scheltens, F. (2010). *Wetenschappelijke verantwoording van de toetsen LOVS rekenen-wiskunde voor groep 3 tot en met 8. [Scientific justification for Mathematics tests for grade 1 to 6]*. Cito.

Jõgi, A. L., Kikas, E., Lerkkanen, M. K., & Mägi, K. (2015). Cross-lagged relations between math-related interest, performance goals and skills in groups of children with different general abilities. *Learning and Individual Differences, 39*, 105–113. https://doi.org/10.1016/j.lindif.2015.03.018.

Kosh, A. E. (2016). *The effects on mathematics performance of personalizing word problems to students' interests* (Unpublished doctoral dissertation). University of North Carolina at Chapel Hill.

Ku, H.-Y., & Sullivan, H. J. (2000). Personalization of mathematics word problems in Taiwan. *Educational Technology Research and Development, 48*(3), 49–59. https://doi.org/10.1007/BF02319857.

Ku, H.-Y., & Sullivan, H. J. (2002). Student performance and attitudes using personalized mathematics instruction. *Educational Technology Research and Development, 50*(1), 21–34. https://doi.org/10.1007/BF02504959.

Logan, S., Medford, E., & Hughes, N. (2011). The importance of intrinsic motivation for high and low ability readers' reading comprehension performance. *Learning and Individual Differences, 21*(1), 124–128. https://doi.org/10.1016/j.lindif.2010.09.011.

López, C. L., & Sullivan, H. J. (1992). Effect of personalization of instructional context on the achievement and attitudes of hispanic students. *Educational Technology Research and Development, 40*(4), 5–14. https://doi.org/10.1007/BF02296895.

Mayer, R. E., Fennell, S., Farmer, L., & Campbell, J. (2004). A personalization effect in multimedia learning: Students learn better when words are in conversational style rather than formal style. *Journal of Educational Psychology, 96*(2), 389–395. https://doi.org/10.1037/0022-0663.96.2.389.

McLaren, B. M., Lim, S. J., Gagnon, F., Yaron, D., & Koedinger, K. R. (2006). *Studying the effects of personalized language and worked examples in the context of a web-based intelligent tutor. Lecture Notes in Computer Science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics), 4053 LNCS* (pp. 318–328). https://doi.org/10.1007/11774303_32.

Moreno, R., & Mayer, R. E. (2000). Engaging students in active learning: The case for personalized multimedia messages. *Journal of Educational Psychology, 92*(4), 724–733. https://doi.org/10.1037/0022-0663.92.4.724.

Murayama, K., Pekrun, R., Lichtenfeld, S., & vom Hofe, R.. (2013). Predicting long-term growth in students' mathematics achievement: The unique contributions of motivation and cognitive strategies. *Child Development, 84*(4), 1475–1490. https://doi.org/10.1111/cdev.12036.

Nelson, E. E., Jarcho, J. M., & Guyer, A. E. (2016). Developmental cognitive neuroscience social re-orientation and brain development: An expanded and updated view. *Accident Analysis and Prevention, 17*, 118–127. https://doi.org/10.1016/j.dcn.2015.12.008.

Niemivirta, M., & Tapola, A. (2007). Self-efficacy, interest, and task performance within-task changes, mutual relationships, and predictive effects. *Zeitschrift Fur Padagogische Psychologie, 21*(3–4), 241–250. https://doi.org/10.1024/1010-0652.21.3.241.

Paas, F. G., van Merriënboer, J. J., & Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills, 79*(1 Pt 2), 419–430.

Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84*(4), 429–434. https://doi.org/10.1037/0022-0663.84.4.429.

Parsons, S., & Bynner, J. (2006). *Does numeracy matter more? National research and development centre for adult literacy and numeracy*. Institute of Education. www.ioe.ac.uk/bedfordgroup.

Prast, E. J., van de Weijer-Bergsma, E., Miočević, M., Kroesbergen, E. H., & van Luit, J. E. H. (2018). Relations between mathematics achievement and motivation in students of diverse achievement levels. *Contemporary Educational Psychology, 55*(August), 84–96. https://doi.org/10.1016/j.cedpsych.2018.08.002.

Ross, S. M., & Anand, P. G. (1987). A computer-based strategy for personalizing verbal problems in teaching mathematics. *Educational Communication and Technology Journal, 35*(3), 151–162. https://doi.org/10.1007/BF02793843.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68.

Sui, J., & Humphreys, G. W. (2015). The integrative self: How self-reference integrates perception and memory. *Trends in Cognitive Sciences, 19*(12), 719–728. https://doi.org/10.1016/j.tics.2015.08.015.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*, 295–312. https://doi.org/10.1016/0959-4752(94)90003-5.

van de Weijer-Bergsma, E., Kroesbergen, E. H., Jolani, S., & van Luit, J. E. H. (2016). Validation of a computerized verbal working memory task for self-reliant administration in primary school children. *Behavior Research Method, 48*(2). https://doi.org/10.3758/s13428-015-0607-y.

van de Weijer-Bergsma, E., Kroesbergen, E. H., Prast, E. J., & van Luit, J. E. H. (2015). Validity and reliability of an online visual-spatial working memory task for self-reliant administration in school-aged children. *Behavior Research Methods, 47*(2), 708–719. https://doi.org/10.3758/s13428-014-0469-8.

Vilenius-Tuohimaa, P. M., Aunola, K., & Nurmi, J. E. (2008). The association between mathematical word problems and reading comprehension. *Educational Psychology, 28*(4), 409–426. https://doi.org/10.1080/01443410701708228.

Walkington, C., & Bernacki, M. (2014). Motivating students by "personalizing" learning around individual interests: A consideration of theory, design, and implementation issues. *Motivational Interventions, 18*, 139–176.

Walkington, C., & Bernacki, M. L. (2018). Personalization of instruction: Design dimensions and implications for cognition. *Journal of Experimental Education, 86*(1), 50–68. https://doi.org/10.1080/00220973.2017.1380590.

Walkington, C., Petrosino, A., & Sherman, M. (2013). Supporting algebraic reasoning through personalized story scenarios: How situational understanding mediates performance. *Mathematical Thinking and Learning, 15*(2), 89–120. https://doi.org/10.1080/10986065.2013.770717.

Weekers, A., Groenen, I., Kleintjes, F., & Feenstra, H. (2011). *Wetenschappelijke verantwoording papieren toetsen Begrijpend Lezen voor groep 7 en 8. [Scientific justification for Reading comprehension test grade 5 and 6]*. Cito.