

# Appraisal of mathematics tasks—a point of view of Norwegian eighth-grade students

Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/27527263241229234  
journals.sagepub.com/home/nea

Ylva Høgset<sup>1</sup>, Marja Van den Heuvel-Panhuizen<sup>1,2</sup> ,  
and Knut Berg<sup>3</sup>

## Abstract

In the large body of literature that is available about the role of motivation in mathematics education, the focus is mostly on students' attitudes and feelings toward mathematics in general. In the current small-scale explorative study, we zoomed in on what students think about the tasks they can come across in mathematics education. To investigate students' appraisal of mathematics tasks, that is, their judgment of how much they like mathematics tasks, we used a mixed method approach with an online questionnaire including closed and open questions. We asked 67 Norwegian eight graders to give an appraisal score for 24 tasks and to express in their own words their reasons for liking/disliking them. In addition, the students also had to indicate for each task whether they think they can solve the tasks. The analysis of the data indicated that the students may prefer bare number problems over context problems. Similarly, students inclined to like puzzle-like tasks more than estimation or straightforward tasks. The reasons for liking/disliking a task most often refer to the difficulty or easiness of the task. The data about the perceived solvability revealed that the more the students believed that they can solve the task, the more they liked the task. Given the different appraisals that the students granted to the different task types and the comments they came with, there is much reason to afford students a stronger voice in mathematics education research, particularly when it is about the tasks used in mathematics education.

## Keywords

appraisal of tasks, perceived solvability of tasks, mathematics tasks, motivation, students

Date received: 15 November 2023; accepted: 4 January 2024

---

<sup>1</sup>Faculty of Education and Arts, Nord University, BODØ, Norway

<sup>2</sup>Freudenthal Institute, Utrecht University, UTRECHT, The Netherlands

<sup>3</sup>Nord University Business School, Nord University, BODØ, Norway

## Corresponding Author:

Marja Van den Heuvel-Panhuizen, Faculty of Education and Arts, Nord University, and Utrecht University, Princetonplein 5, 3584 CC, Utrecht, The Netherlands.

Email: m.vandenheuvel-panhuizen@nord.no

## 1. Introduction

It has been revealed in several studies (e.g., Forsblom et al., 2022; Schukajlow, 2015) that students' feelings about mathematics, their emotions, and motivational orientations such as enjoyment, boredom, and interest are of influence on their learning of mathematics and are related to the achievement level they acquire in mathematics. For example, in the recently carried out longitudinal large-scale survey by Forsblom et al. (2022) in which the mathematics-related emotions of 10–17-year-old students were investigated, it was found that enjoyment of mathematics positively predicted students' achievements, while anger and boredom negatively predicted their learning results.

However, we noticed that when investigating students' motivation for mathematics, students' liking/disliking of tasks is scarcely researched. The work of Schukajlow (2015) is an exception to this. He scrutinized the relation between boredom and performance and particularly examined this for three types of problems. In his experimental study with ninth and tenth graders, he found significant differences in how a relation between boredom and performance worked out for different types of problems. The correlations tended to be lower for intramathematical problems, which do not have a connection to reality and can be solved by using only mathematical procedures than for word problems or modeling problems.

This finding makes it clear that when investigating students' liking or disliking of mathematics, the nature of the tasks is a factor that should be taken into account. Inspired by the research of Schukajlow (2015) our study continued on this line of inquiry.

## 2. Theoretical perspective and research questions

The theoretical perspectives that guided our study were the crucial role of motivation in learning mathematics and the specific role of tasks to promote interest and engagement in students. Together with the findings from earlier research on students' appraisal of tasks, these perspectives formed the basis of our research questions.

### 2.1 Crucial role of motivation

Motivation is considered to be the driving force in the learning of mathematics (e.g., Gojak, 2013). It makes sure that students are actively involved in learning. Their positive feelings may impact their willingness to be involved and increase their motivation. On the other hand, negative feelings may result in a decrease in their motivation. Therefore, a great deal of research has been carried out aimed at getting a better understanding of students' motivation.

Concerning motivation in the context of learning mathematics an important basis for this research is led by McLeod (1992) who conceptualizes affect as consisting of emotions, attitudes, and beliefs. Following this approach, Hannula (2006, p. 166) sees motivation as "a potential to direct behaviour that is built into the system that controls emotion." When it is about the motivation to solve a mathematics task, Hannula (2006) conceives that students' motivation to problem-solving as manifested in their beliefs about their cognition, behavior, and emotion.

Middleton et al. (2017, p. 668) see "affect, cognition, motivation, and social interactions as intertwined, mutually interacting, and highly situational influences on engagement." Here engagement means that somebody takes part in an activity that requires some cognitive or affective investment. Middleton et al. (2017) refer to this as in-the-moment engagement which can be connected to a very specific object such as a mathematics task. According to Mitchell (1993), this interest in the moment needs to be caught by some key facets of the task. This may mean that students' interest in a task is not only an issue of their levels of personal interest in mathematics. Since, regardless of students' general feelings about the subject the interest they show in any particular task can vary widely (Middleton et al., 2017).

The approach of taking into account the situational influences can also be found in the work of Schukajlow et al. (2023), who recognized that there is recently increased attention on the situational nature of affective constructs. Therefore, they decided to include in their survey paper, “situational specificity” as a key characteristic of emotions and motivation which can be embedded in social environments and sociocultural contexts as well as in tasks. For example, researchers found that adapting the contexts of tasks to students’ interest areas improved the students’ situational interest in the task (Bernacki & Walkington, 2018) and that prompting students to find multiple solutions to modeling problems had a positive influence on students’ interest in mathematics (Schukajlow & Krug, 2014).

Such a task-specific approach is scarcely found in motivation studies. The focus is mostly on mathematics in general or on what students think of parts of mathematics. Specifically, in TIMSS the general question of liking/disliking mathematics is one of the standing components in the student questionnaire. The 2019 cycle of this study covering 64 countries figured out that only 20% of the eighth-grade students liked learning mathematics very much, 39% liked it somewhat, and 41% did not like it (Mullis et al., 2020).

Given the consequences of this negative attitude toward mathematics, the question immediately arises as to what the ground is for students not liking mathematics. When the focus is on mathematics in general then the provided reasons for liking/disliking will also be general. For example, when Aguilar (2021) asked high school students about this he found that one of the main reasons for student’s reluctance toward mathematics is their perceived lack of understanding and their self-perception of low content knowledge. Comparable reasons were found in a study in Tanzania carried out by Mazana et al. (2019), in which a large group of primary, secondary, and college students participated. The students mentioned that they do not like mathematics because the subject is difficult, they are not good at mathematics, and it is not necessary for their future careers. Students who liked mathematics indicated that they like it because they see its value in everyday life, and because they find mathematics interesting and enjoyable.

Studies like those previously mentioned provide information about how students think about mathematics in general, but they do not offer concrete indications for how education can be adapted to facilitate students’ engagement.

## 2.2 *Specific role of tasks*

To find ways for improving students’ engagement requires a deeper look into where the mathematical activity starts and where the motivation is triggered. According to Middleton (2013, p. 78), it is the task where the motivation is induced, “because much of students’ motivation in mathematics is due to the way we design mathematical tasks.” Similarly, Roche et al. (2021) also identified the nature of a task and whether it is intrinsically interesting for students as relevant for increasing the likelihood of their engagement. Tasks that students like can act as facilitators to make them more motivated to work with them and can create advantageous conditions to make the learning process successful.

This pivotal role of the tasks that are used in mathematics education has also been emphasized by several other scholars. Among other things, they came to the conclusions that mathematical tasks largely determine what students learn (Hiebert & Wearne, 1993), that tasks have the greatest influence on what students learn compared to other features in the classroom (Smith & Stein, 1998), that tasks can shape students’ mathematical learning opportunities (Watson & Mason, 2007) and that the chosen tasks have a major part in determining the quality and nature of their learning (Clarke & Roche, 2018).

## 2.3 *Earlier research on students’ appraisal of tasks*

Despite the relevance that is attributed to tasks, research on how students think about different tasks is scarce. The study of Sullivan et al. (2013) is an exception. They chose to focus their data collection on

finding insights into students' reactions to different types of tasks since task characteristics seemed to be the main determinant of students' decisions on engagement. In the survey, in which 922 students from Grades 5 to 8 were asked to rank three types of tasks, it was found that 54% chose the number calculation as their first preference. For the word problem, this was 35%, and for the open-ended task, this was only 12%. For preferring number calculation and word problems, the most often-given reason was finding them to be easy. For putting open-ended tasks in the first place the most frequently mentioned reasons referred to the challenging character of the task: "I like it because I want something challenging and hard."

In a further analysis, Sullivan et al. (2013) found that the group of students who preferred number calculation contained more students who felt confident in mathematics than students with low confidence. Maybe the students with the highest confidence might prefer the number calculation task because then their success is under their own control. For the word problem, the reverse was the case. Then there were more students in the lowest confidence than in the highest confidence category. For the open-ended task, there was almost no difference between the numbers of students in the one and the other category of confidence.

These findings, and particularly the finding that students who had the highest confidence put the number calculation on the first place because of the easiness of the task, make clear that the students' confidence with respect to whether they think they can solve the task, that is, the perceived solvability, is a relevant factor when investigating the students liking of the tasks. This is in line with Bandura's (1997) theory about self-efficacy which refers to the judgments individuals have about their perceived capability to perform future tasks and of which the perceived level of task difficulty is an integral part (Street et al., 2022). According to Bandura (1993, p. 119), "positive attitudes toward mathematics were better predicted by perceived self-efficacy than by actual ability." This means that mathematics self-efficacy influences students' motivational orientation. Students who report higher self-efficacy also report greater intrinsic motivation.

Support for this is furthermore provided by Stevens et al. (2004). They found in ninth- and 10th-grade students that self-efficacy predicted their motivational orientation and subsequently their mathematics performance. Also, in several other studies indications can be found about the relation between mathematics self-efficacy and the attitude toward mathematics (Öztürk et al., 2020) and the finding that one of the main reasons for student's reluctance toward mathematics is their perceived lack of understanding and their self-perception of low content knowledge (Aguilar, 2021), and that "student motivation – both interest and motivated behaviour – is strongly predicted by self-efficacy" (Skaalvik et al., 2015, p. 135).

The basis for our study was what these earlier studies have shown about the possible influence of the task type, the reasons for liking/disliking tasks and the role of self-efficacy, especially understood as how students think about whether they can solve the tasks. Therefore, we formulated the following research questions:

1. What types of mathematics tasks do students tend to like or dislike?
2. Why do students like or dislike these tasks?
3. What is the relation between students' appraisal of tasks and the perceived solvability of the tasks?

### **3. Method**

#### ***3.1 Set up of the study***

To answer the three research questions, a survey was carried out among eighth-grade students in Norway aimed at collecting data about the students' appraisal of different kinds of mathematics tasks and about their reasons for liking or not liking the tasks. To know whether the students'

appraisal of the tasks was related to their perceived solvability of the tasks it was also investigated whether the students thought they could solve the tasks in question.

**Participants.** For recruiting students to take part in the study, an email was sent to the principals of about 20 schools in Northern Norway with information about the project and an invitation to participate in the study with their eighth-grade mathematics classes. The contacted schools were both large and small size, covering either Grades 1–10 or Grades 8–10. In the end, two large and two small schools in the Nordland region accepted to participate which resulted in a sample of 67 eighth-grade students. The large schools had multiple classes per grade, but less than 500 students in total. The small schools had around 100 students or less and only one class per grade.

### 3.2 Instrument

The instrument used for this study is an anonymous online questionnaire which was implemented in nettskjema.no. The questionnaire consisted of 24 mathematics tasks mainly in the domain of numbers for which each three questions were asked. First, the students had to indicate for each task on an ordinal Likert scale whether they like or dislike the task. There were five options for answering, ranging from liking the task a lot to strongly disliking it. Second, they were asked to explain in their own words why they like or dislike the task. The third question to be answered was whether they think they can solve the task. It was made clear to the students that for answering this question it was not necessary to solve the task. For this third question also an ordinal Likert scale with five options available was used which now ranged from being very certain they can solve the task to being very certain they cannot solve it. Before putting the final version of the questionnaire online, a try-out was carried out to check whether the tasks and the questions were clearly formulated, and whether the number of tasks was feasible for the students.

Each task was visible on the screen together with the three questions (Figure 1), which made that the students could see the task when they were answering the questions. At the start of the questionnaire, information was included to explain the questionnaires to the students and at the end, there was a message with a “thank you” for answering the questionnaire.

The students received the link to the questionnaire from their teacher. No further requirements were necessary for logging in, and no personal information had to be put in. The questionnaire was completely anonymous. From November 25, 2021 to January 15, 2022 the questionnaire was open for answering. The students all responded in early January 2022.

The tasks included in the questionnaire were initially taken from two mathematics textbooks for Grade eight. Moreover, to get a somewhat larger variety of tasks a few of them were altered or replaced by some self-made tasks. The used textbooks were published recently, and are following the new curriculum for mathematics education in Norway (Kunnskapsdepartementet, 2020). In order to limit the effect that the difficulty level of the tasks could have on the students' appraisal, tasks were chosen which were marked in the textbooks as being of medium difficulty. Furthermore, the tasks were mostly taken from the first chapters of the textbooks so that it would be likely that the students would have already learned about them. In this way, unfamiliarity with the topics addressed in the questionnaire would have less influence on the answers of the students.

Table 1 gives an overview of the task types included in the questionnaire. The tasks were classified on how the problem is presented (as a bare number problem or as a context problem), the nature of the problem requiring a particular way of mathematical processing (straightforward calculating, puzzling, or estimating), the operation involved ( $+$ ,  $-$ ,  $\times$ ,  $\div$ ) and the number types involved in the task (whole numbers, fractions, decimals, percentages). For every subcategory of task type, Table 1 shows how many of these were in the questionnaire.

**Oppgave 1**

$347 + 489 =$

a. Hva synes du om denne oppgaven?

Jeg liker den veldig godt

Jeg liker den litt

Jeg verken liker eller misliker den

Jeg misliker den litt

Jeg misliker den sterkt

b. Hvorfor liker/misliker du denne oppgaven?

c. Tror du at du ville klart å løse oppgaven?  
(Du trenger ikke å faktisk løse den.)

Ja, jeg er veldig sikker på at jeg kan løse den

Ja, men jeg er bare litt sikker på at jeg kan løse den

Jeg vet ikke om jeg kan løse den

Nei, jeg tror jeg mest sannsynlig ikke kan løse den

Nei, jeg er veldig sikker på at jeg ikke kan løse den

a. What do you think of this task?

I like it very much

I like it a little

I neither like nor dislike it

I dislike it a little

I strongly dislike it

b. Why do you like or dislike this task?

c. Do you think that you would be able to solve this task?  
(You do not need to solve it.)

Yes, I am very sure that I can solve it

Yes, but I am only a bit sure about it

I do not know whether I can solve it

No, I think it is most likely that I cannot solve it

No, I am quite sure that I cannot solve it

**Figure 1.** Screen view from the questionnaire and the English translation.

**Table 1.** Type and number of tasks included in the questionnaire.

Task type	Subcategory	Number of tasks
Presentation of the problem	Bare number problem	10
	Context problem	14
Nature of the problem	Straightforward problem	14
	Puzzle-like problem	7
	Estimation problem	3
Operation involved	+	2
	-	3
	×	2
	÷	1
	Mixed	16
Number type involved	Whole numbers	16
	Fractions	4
	Decimals	2
	Percentages	2

As can be seen in Table 1, the number of tasks is not the same for every subcategory. When choosing the number of tasks, it was attempted to follow the proportion that students can come across these categories regularly in textbooks. The only exception made was for context problems. In textbooks, the proportion of context problems is normally lower than that of bare number problems. The reason for including more context problems was to make it possible to investigate whether the appraisal of context problems and bare number problems differ. When calculating the Cronbach's alpha of the data collected for the 24 tasks, we found sufficient reliability for both the appraisal of tasks ( $\alpha = .957$ ) and the solvability of the tasks ( $\alpha = .965$ ).

The tasks were ordered in the questionnaire in such a way that there were not too many of a certain type after each other. This was done so that the order in which the students would see the tasks would have less of an effect on how they would answer the questions. The first few tasks were among the easier ones. This was done in order not to make the students lose their motivation in the beginning.

### 3.3 Data analysis

To investigate what tasks students like (Research Question 1) and whether students' perception about being able to solve a task was related to liking or not liking a task (Research Question 3) a quantitative analysis was carried out. The analysis of the reasons for liking or disliking a task (Research Question 2) was mainly done more qualitatively by identifying categories of reasons and structuring them.

The responses the students gave to the first and third questions in the questionnaire (appraisal score and solvability score, respectively) were analyzed by means of IBM SPSS Statistics version 27. To carry out this analysis, the scores on the two rating scales were transposed to the numbers 1 to 5 where 1 was the lowest appraisal and perceived solvability, and 5 was the highest. First, the mean appraisal score and the standard deviation of each individual task were calculated. Since we measured the appraisal and solvability score at an ordinal level, we also calculated the medians of the scores. Thereafter the same calculations were carried out for each type of task. This was followed by testing whether there were differences between the different types. To find out whether there was a relation between the perceived solvability of the tasks and the appraisal score, Pearson's correlation coefficient was calculated on the basis of all the cases in which a score for solvability and appraisal was given to a task.

The reasons given for liking or disliking a task were analyzed by first importing all the written responses into an Excel file in which for each task and for each appraisal score the responses were sorted. In the second step, the given reasons were clustered in a number of main categories (Table 2). This clustering was first done by the first author, which was followed by a check by the second author and a joint discussion to reach an agreement. Some of the categories we found are similar to the ones that appeared in previous literature such as Mazana et al. (2019). After establishing the categories of reason, the frequency of each was calculated per appraisal score for each task and overall the tasks together.

## 4 Results

### 4.1 Students' appraisal of mathematics tasks

Out of the 1,608 possible ratings that the students could provide, 40 were left blank. This means that for the 24 tasks 1,568 ratings from 1 to 5 were given in total by 67 students. The analysis of the data indicated that in general the students neither liked the tasks very much nor disliked them very much ( $M = 3.25$ ;  $Mdn = 3$ ;  $SD = 1.255$ ). Table 3 contains the mean scores and standard deviations of the students' appraisal for each task. The tasks which were liked the most are colored blue and those

**Table 2.** Main categories of reasons for liking or disliking the tasks.

Main category	Description
Liking	Includes saying that they like or love the task, and saying that the task is good, nice, ok, or fine. Does not include specifying what task they like.
Disliking	Includes saying that they dislike the task, the type of task, or mathematics in general. Also includes calling the task annoying or irritating. Does not include specifying what of the task they dislike.
Easiness	Includes saying the task is easy/simple, that it is too easy, saying they understand/can solve the task, or that it is not difficult.
Difficulty	Includes saying that the task is difficult, that they cannot solve it, that they do not understand the task, that they believe themselves to not be smart enough to solve the task, that the explanation is bad, that there is not enough information, that it takes a long time, that it is confusing or complicated, that there is too much information or that the numbers are too large.
Moderateness	Includes saying that the task is not too difficult or too easy, and saying that it is difficult but not too difficult.
Challenge	Includes saying that you have to think or puzzle when solving the task. Also saying that one has to spend time on the task, put effort into solving it, or that they have to explain the answer. There are some similarities with difficulty, but the difficulties are lesser and more manageable.
Fun	Includes saying the task is fun, interesting, creative, satisfying to solve, different, or not boring.
Boring	Includes calling the task boring or repetitive.
Mathematical content	Includes referring to various mathematical content domains and topics.
Context	Includes mentioning the specific situation in which the task is embedded.
Text	Includes mentioning that there is text in the task, or that one has to write.
Usefulness	Includes mentioning how useful, useless or practical the task is.
Familiar	Includes saying the task is familiar, common, normal, or basic.
Don't know	Includes saying that one is unsure or does not know why the task is liked or disliked, or saying that it is difficult to explain.
Other reason	Includes responses that are not related to the question, that use emojis instead of words, or that are not fitting into another category.

which were disliked the most are colored orange. The specific tasks used in this study are in Table 3b in the Appendix.

The tasks that the students liked the most were three tasks with straightforward bare number problems (Tasks 1, 2, 14), two tasks with a puzzle-like bare number problem (Tasks 5 and 16), one task with a straightforward problem in context (Task 20), and one task with a puzzle-like problem in context (Task 7). All these tasks got a score of 5 from 20% to almost 40% of the students.

The less liked tasks were two tasks about fractions in context (Tasks 3 and 22), one about fractions with bare numbers (Task 8), one about percentages in context (Task 9), and one with a puzzle-like problem in context (Task 13). All these tasks had a mean appraisal score lower than 3.0. Of the complete set of tasks, an appraisal score of 1 was given most often to Tasks 8 and 9. This was done by 22.7% and 19.4% of the students, respectively. Task 9 is a context task about percentages and Task 8 is a bare number problem about fractions. This task also had the lowest average appraisal score.

**4.1.1 Students' appraisal of bare number tasks and context tasks.** In the appraisal scores for the subcategories of TaskType1 (tasks with bare number problems or tasks with context problems), the highest appraisal scores were found for the tasks including bare number problems. The highly appraised tasks



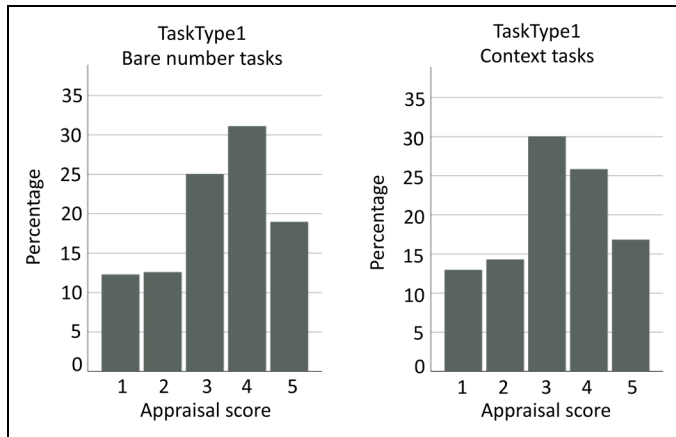
**Table 3.** Mean scores and standard deviation for the students' appraisal of tasks.

		Mean score	SD	Absolute and relative frequency				
				Score 1	Score 2	Score 3	Score 4	Score 5
1	+ bare numbers	3.7879	1.08861	4 6.1%	3 4.5%	14 21.2%	27 40.9%	18 27.3%
2	- bare numbers	3.5522	1.22206	6 9%	6 9%	17 25.4%	21 31.3%	17 25.4%
3	Fractions in context	2.9701	1.15431	7 10.4%	18 26.9%	18 26.9%	18 26.9%	6 9%
4	Investigation with estimation in context	3.0625	1.31987	10 15.6%	11 17.2%	20 31.3%	11 17.2%	12 18.8%
5	Puzzle-like problem bare numbers	3.7164	1.02361	6 9%	4 6%	12 17.9%	26 38.8%	19 28.4%
6	+/ $\times$ in context	3.0909	1.19907	10 15.2%	8 12.1%	20 30.3%	22 33.3%	6 9.1%
7	Puzzle-like problem in context	3.5909	1.16325	5 7.6%	5 7.6%	18 27.3%	22 33.3%	16 24.2%
8	Fractions bare numbers	2.4848	1.09884	15 22.7%	18 27.3%	21 31.8%	10 15.2%	2 3%
9	Percentages in context	2.9104	1.27602	13 19.4%	11 16.4%	19 28.4%	17 25.4%	7 10.4%
10	Puzzle-like problem in context	3.3333	1.37375	10 15.2%	8 12.1%	14 21.2%	18 27.3%	16 24.2%
11	Percentages bare numbers	3.2273	1.28656	10 15.2%	8 12.1%	15 22.7%	23 34.8%	10 15.2%
12	Decimals in context	3.1563	1.25	8 12.5%	11 17.2%	18 28.1%	17 26.6%	10 15.6%
13	Puzzle-like problem in context	2.9841	1.07	6 9.5%	13 20.6%	25 39.7%	14 22.2%	5 7.9%
14	$\times$ bare numbers	3.5909	1.30062	6 9.1%	9 13.6%	11 16.7%	20 30.3%	20 30.3%
15	Decimals in context	3.2615	1.22827	8 12.3%	7 10.8%	21 32.3%	18 27.7%	11 16.9%
16	Puzzle-like problem bare numbers	3.6	1.18322	7 10.8%	3 4.6%	12 18.5%	30 40.2%	13 20%
17	Fractions bare numbers	3.0154	1.20536	9 13.8%	12 18.5%	20 30.8%	17 26.2%	7 10.8%
18	Investigation with estimation in context	3.3016	1.19993	7 11.1%	6 9.5%	22 34.9%	17 27%	11 17.5%
19	Puzzle-like problem bare numbers	3.0455	1.20807	9 13.6%	11 16.7%	22 33.3%	16 24.2%	8 12.1%
20	+/ $\times$ in context	3.7846	1.25614	5 7.7%	5 7.7%	14 21.5%	16 24.6%	25 38.5%
21	$\div$ bare numbers	3.1538	1.26529	9 13.8%	9 13.8%	21 32.3%	15 23.1%	11 16.9%
22	Fractions in context	2.8	1.09259	9 13.8%	14 21.5%	28 43.1%	9 13.8%	5 7.7%

(continued)

**Table 3.** (continued)

				Absolute and relative frequency				
		Mean score	SD	Score 1	Score 2	Score 3	Score 4	Score 5
23	Puzzle-like problem in context	3.1077	1.27626	11 16.9%	7 10.8%	20 30.8%	18 27.7%	9 13.8%
24	Investigation with estimation in context	3.3492	1.32176	9 14.3%	6 9.5%	16 25.4%	18 28.6%	14 22.2%
		3.2455	1.25455	199 12.7%	213 13.6%	438 27.9%	440 28.1%	278 17.7%



**Figure 2.** Relative frequency distributions in percentages of the given appraisal scores for tasks with a bare number problem ( $n = 659$ ) and tasks with a context problem ( $n = 909$ ).

with a bare number problem were three straightforward calculation problems and two puzzle-like problems. Of the 10 bare number problems, five had an average score higher than 3.5, and of the 14 tasks with a context problem, there were only two with an average score higher than 3.5. These two tasks are both related to playing football. Four of the tasks with a context problem had a mean score of less than 3.0, while this was only the case for one task with a bare number problem.

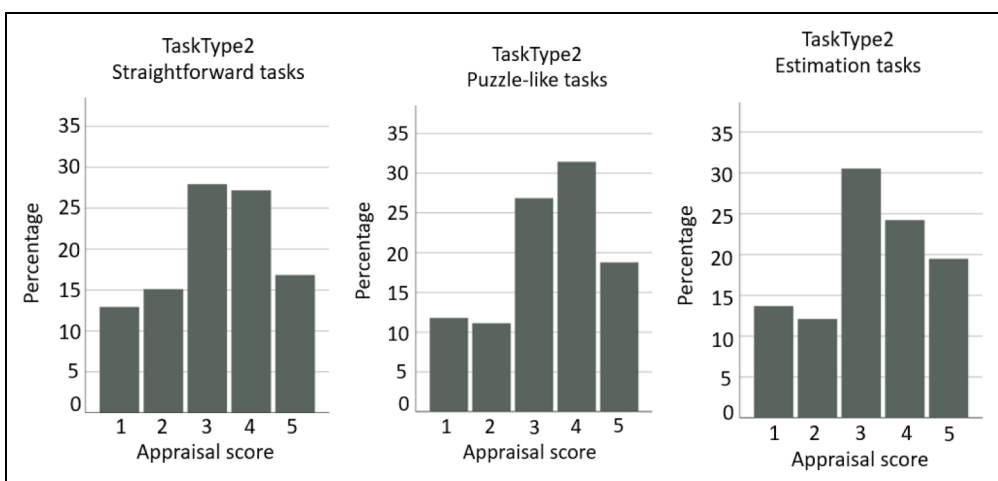
Altogether the tasks with a bare number problem had a higher mean, median, and standard deviation than the tasks with a context problem (bare number problems:  $M = 3.32$ ;  $Mdn = 4$ ;  $SD = 1.260$ ; context problems:  $M = 3.19$ ;  $Mdn = 3$ ;  $SD = 1.248$ ). As shown in Figure 2, for the tasks with a bare number problem the most often appraisal score given was 4, and the score given second most often was 3. The reverse was the case for the tasks with the context problems, for which the score 3 was given the most often and the score 4 the second most often. Furthermore, the bar graphs also show that the tasks with bare number problems have a slightly lower percentage of disliking the tasks and a slightly higher percentage of liking the tasks than the tasks with context problems.

The above findings tend to indicate that the bare number tasks were liked more than the context tasks. However, by using a Wilcoxon matched-pair signed rank test and computing for the two categories the averages of each student's appraisal scores yielded no significant results ( $n = 67$ ,  $T = 766$ ,  $p = .067$ ).

Finding no significant difference between the two categories corresponds with our finding that the tasks that were liked most were not always bare number tasks. Of the seven tasks with the highest appraisal score, five were bare number tasks and two were context tasks. In addition, the task that was disliked the most was not a context problem but a bare number task about fractions. This shows that there are more task characteristics that may influence students' appraisal of a mathematics task.

**4.1.2 Students' appraisal of straightforward, puzzle-like, and estimation tasks.** Students' appraisal scores for the three subcategories of TaskType2 (tasks with straightforward problems, tasks with puzzle-like problems, and tasks with estimation problems) were largely in accordance with each other. The descriptive statistics of the tasks with straightforward problems and the tasks with estimation problems were about the same (straightforward problems:  $M=3.20$ ;  $Mdn=3$ ;  $SD=1.255$ ; estimation problems:  $M=3.24$ ;  $Mdn=3$ ;  $SD=1.281$ ). The mean and median of the tasks with puzzle-like problems were a bit higher ( $M=3.34$ ;  $Mdn=4$ ;  $SD=1.239$ ). The standard deviation was the highest for the estimation problems and the lowest for puzzle-like problems.

As is shown in Figure 3, some differences could be found in the frequency distribution of the given appraisal scores. Tasks with puzzle-like problems got a slightly lower percentage of students who gave an appraisal score of 1 and 2 than tasks with straightforward and estimation problems. Furthermore, tasks with estimation problems obtained a slightly higher percentage of students who gave an appraisal score of 5. In sum, among the different natures of the tasks within TaskType2 there was not one type of task that was clearly preferred in all ways, but it seems that the tasks with puzzle-like problems were the closest to this state. Computing the averages of each student's appraisal scores for the three categories of tasks yielded no significant results by using Friedman's two-way ANOVA ( $n=67$ ,  $Q=4.635$ ,  $p=.099$ ). However, when comparing the straightforward tasks with the puzzle-like tasks, a Wilcoxon matched-pair signed rank test showed a significant difference in appraisal of the advantage of the puzzle-like tasks ( $n=67$ ,  $T=1355$ ,  $p=.017$ ).



**Figure 3.** Relative frequency distributions in percentages of the given appraisal scores for tasks with a straightforward problem ( $n=920$ ), tasks with a puzzle-like problem ( $n=458$ ), and tasks with an estimation problem ( $n=190$ ).

**4.1.3 Relations between task types and students' appraisal of the tasks.** To investigate whether the presentation type of a task (bare number and context) and the nature type of a task (straightforward and puzzle-like) matters for whether the students like/dislike a task, we took a closer look at how the appraisal scores between the subcategories of TaskType1 and TaskType2 were related. Estimation problems were left out in this analysis because we only had estimation problems in a context format and not in a bare number format.

Although comparing bare number tasks and context tasks did not make a significant difference as shown earlier, this might change when we bring in the nature of the problem. For tasks with a straightforward problem, a Wilcoxon matched-pair signed rank test revealed that there was still not a significant difference between bare number tasks and context tasks ( $n=67$ ,  $T=603.5$ ,  $p=.076$ ). However, for the puzzle-like tasks, this was different. The appraisals for the puzzle-like bare number tasks were significantly higher than for the puzzle-like context tasks ( $n=67$ ,  $T=577.5$ ,  $p=.020$ ). Furthermore, we found a significantly higher appraisal for the bare number tasks with a puzzle-like problem than for the bare number tasks with a straightforward problem ( $n=67$ ,  $T=1361$ ,  $p=.016$ ).

## 4.2 Students' reasons for liking or disliking mathematics tasks

As is shown in Table 4 (Part A), in 930 (58%) of the 1,608 cases (67 students each with 24 tasks) in which the students could give their opinion about the tasks, a reason was provided for liking or disliking the task.

Part of the 678 cases in which the reason was missing was because in Task 14, the question about the reason was unintendedly left out. There were also nine cases in which the students gave a reason but no appraisal score, and in 31 cases, both were not provided. Table 4 (Part B) gives each appraisal score the absolute and relative frequency of the main categories of reasons that were mentioned for liking or disliking the tasks.

## 4.3 Specific reasons for liking or disliking mathematics tasks

**4.3.1 Reasons related to the difficulty or easiness of the task.** The reasons that were mentioned most often for liking or disliking tasks were those in which something was said about the difficulty or easiness of the task. Difficulty was mentioned in 20.6% of the cases and for easiness, this was 18%. A task being difficult was the most often given reason for an appraisal score of 1 and 2 and a task being easy was the most often given reason for an appraisal score of 4 and 5.

In seven cases for which a score of 1 or 2 was given, the students said that they did not know how to solve the task. Then their reaction was for example, "don't know how to do it" and "I don't know where I would start to solve it." In 16 cases, students referred to their own perceived lack of skill or intelligence, like in the cases where they said, "I'm bad at this" and "because my head is not smart at all." In addition, in 29 of the answers, it was explicitly indicated that the task was difficult, complicated, or confusing. The students expressed this by saying "I don't like it because it is a bit too difficult," "have never liked this sort of task[; ]very confusing," and "it is complicated." For the appraisal score of 3, the replies were similar to the appraisal scores of 1 and 2, but here the students did not mention lack of skill or intelligence as the reason for their score. Answers were given such as "this is a good task but a little difficult," and "seems a bit advanced." There were also students who gave an appraisal score of 4 and nevertheless said that a task was difficult. Some of these answers were: "I'm not good at these sorts of tasks," "don't quite understand the task," and "Because it is a little complicated."

Saying that a task is easy was most often a reason to like a task. However, the opposite did also happen. In 14 of the cases, it was the easiness of the task that obviously made the students give a low

**Table 4.** Frequency of reasons given and not given per appraisal score (Part A) and frequency of the reasons given per appraisal score (Part B).

Reason	No appraisal score	Appraisal score 5	Appraisal score 4	Appraisal score 3	Appraisal score 2	Appraisal score 1	Total
<b>Part A</b>							
Reason given	9 (22.5%)	148 (53.2%)	240 (54.4%)	221 (50.5%)	153 (71.8%)	159 (79.9%)	930 (57.8%)
No reason given	31 (77.5%)	130 (46.8%)	200 (45.5%)	217 (49.5%)	60 (28.2%)	40 (20.1%)	678 (42.2%)
Total	40 (100%)	278 (100%)	440 (100%)	438 (100%)	213 (100%)	199 (100%)	1608 (100%)
<b>Part B</b>							
Liking	1 (11.1%)	12 (8.1%)	31 (12.9%)	20 (9%)	5 (3.2%)	33 (20.8%)	64 (6.9%)
Disliking			2 (0.8%)	8 (3.6%)	5 (3.2%)	9 (5.7%)	48 (5.2%)
Easiness	3 (33.3%)	60 (40.5%)	72 (30%)	21 (9.5%)	5 (3.2%)	43 (27%)	167 (18%)
Difficulty			14 (5.8%)	74 (33.5%)	58 (37.9%)	43 (27%)	192 (20.6%)
Moderateness			12 (5%)	4 (1.8%)	6 (3.9%)	1 (0.6%)	16 (1.7%)
Challenge	1 (11.1%)	11 (7.4%)	21 (8.8%)	15 (6.8%)	6 (3.9%)	1 (0.6%)	55 (5.9%)
Fun		14 (9.5%)	28 (11.7%)	4 (1.8%)	13 (8.5%)	27 (17%)	46 (4.9%)
Boring			2 (0.8%)	8 (3.6%)	42 (27.5%)	20 (12.6%)	50 (5.4%)
Math content	1 (11.1%)	25 (16.9%)	20 (8.3%)	29 (13.1%)	13 (8.5%)	20 (12.6%)	137 (14.7%)
Context	1 (11.1%)	11 (7.4%)	12 (5%)	1 (0.5%)	42 (27.5%)	1 (0.6%)	26 (2.8%)
Text	1 (11.1%)	3 (2.0%)	3 (1.3%)	6 (2.7%)	14 (9.2%)	4 (2.5%)	31 (3.3%)
Usefulness		4 (2.7%)	3 (1.3%)	1 (0.5%)	1 (0.6%)	1 (0.6%)	8 (0.9%)
Familiar		1 (0.7%)	2 (0.8%)	1 (0.5%)	1 (0.7%)	1 (0.6%)	5 (0.5%)
Don't know		1 (0.7%)	10 (4.2%)	20 (9%)	6 (3.9%)	2 (1.3%)	39 (4.2%)
Other reason	1 (11.1%)	6 (4.1%)	8 (3.3%)	10 (4.5%)	3 (2%)	18 (11.3%)	46 (4.9%)
Total	9 (100%)	148 (100%)	240 (100%)	221 (100%)	153 (100%)	159 (100%)	930 (100%)

appraisal score. One student expressed this as follows: “Because it is way too easy.” Other students often reacted in these cases with only the words “easy” or “simple.” Easiness was also mentioned in 21 of the cases where students gave an appraisal score of 3. Some answers were short by just saying “easy,” and others were a bit longer, such as “way too easy,” “These are not difficult,” and “It is very easy to do this one. You just have to do minus 1.5 on both sides.”

The reasons that refer to easiness when giving an appraisal score of 4 and 5 indicated that the students were able to solve or understand the task. Then responses came up such as “because it is easy,” “because it is easy to find the answer to it,” “I think I can solve this task,” and “not that difficult.”

**4.3.2 Reasons related to the mathematical content of the tasks.** In 137 cases, the students referred to specific mathematical content as a reason for liking or not liking the tasks. This involved 15% of all the cases in which a reason was given. As shown in Table 5, the mathematical topics mentioned most often were operations with whole numbers and fractions. Topics also brought up were percentages and mental calculations. In a few cases, estimation and measurement were referred to.

Although these content-related reasons were found for all five appraisal scores, most often the mathematical content was the reason for giving a low score. In 62 cases, the task was given an appraisal score of 1 or 2. These scores were mostly found when the reason referred to the topic of fractions, percentages, division, or subtraction. Many, but not all, responses were short and did not go into much detail. For example, students wrote “I think fractions are a little difficult” and “don’t like division.” An example of a longer response is “calculating with percent is probably nice to know but I don’t understand it and then it is difficult to like it.” One student showed ambivalence and was not entirely negative and called a task useful despite giving an appraisal score of 2 and reacted with “don’t remember anything about calculating percent but it is a useful task.” When students gave a low appraisal score because of the mathematical content, their reason was related to not liking the content, not understanding it, or both.

For the appraisal score of 3, there were 29 cases in which the reason had a reference to the content. The most common reasons for giving this score were again the topics of fractions, division, percentages, and subtraction. Only a few times were estimation, area, addition, and mental calculation mentioned. A typical reason that was mentioned together with an appraisal score of 3 was “don’t like

**Table 5.** Frequency of mentioning a mathematical topic as a reason for liking or disliking tasks.

Topic	<i>n</i>	%
Percentages	26	19
Operations	42	30.7
+	9	
–	11	
×	1	
	19	
Other <sup>a</sup>	2	
Fractions	32	23.4
Mental calculation	21	15.3
Estimation	8	5.8
Measurement	8	5.8
m/cm	3	
Area	3	
Liter	2	
Total	137	100

<sup>a</sup>Here the missing numbers in the task were mentioned as a reason.

minus, but it would go well enough.” There were a few other responses that also had both a positive and a negative connotation, like “as I said, I don’t like calculating with fractions, but I can do it if I try” and “I’m not that good at multiplying with fractions but I think it is fun to try.” Other responses were more neutral, like “I think the task is completely ok because it is minus.” Some of the responses indicated that the appraisal score of 3 was mainly considered negative. Here reasons were given such as “Don’t like calculating with litres that much” and “I am a little bit bad at percent.”

The content-related reasons that were mentioned in the 20 cases where an appraisal score of 4 was given, were mental calculation, the addition operation, the division operation, and percentage. Topics mentioned less often were the subtraction operation, fractions, meter/centimeter, multiplication, and estimation. Students gave responses like “I like the task a little because I am ok at division,” and “I like fractions a little.” Also, even a more positive response like “calculating with percent is fun” was given. Some students gave reasons such as “I am not good with minus” and “Division is a little confusing,” even though they gave it a score of 4.

In 25 of the cases in which an appraisal score of 5 was given, reasons were mentioned that referred to mental calculation, the addition operation, and percentages (3 times). Also, fractions and estimation were brought up one time each. Many of the responses were short and said that the students liked the task because they liked the mathematical content, or that the content was easy. Students came with responses such as “I like calculating percent” and “easy it is just plus.” A response that went into more detail was: “I like tasks where you have to estimate how much you do something like for example, how much electricity do you use, how many calories do you eat, and those kinds of things.”

Overall, it can be concluded that the results about the mathematical content of the tasks did not play a dominant role in liking/disliking a task. Only in a small proportion of the cases, a reason was given in which there was a reference to the mathematical content. Reasons for not liking a task brought up the well-known stumbling points of mathematics education: fractions, percentages, division, and subtraction. Yet, contrary to what might be expected, we did not really find a division between content that either was liked or was disliked. For example, these four mathematical content domains were also mentioned as a reason for rather liking a task.

**4.3.3 Reasons related to the context of the tasks.** In total, the questionnaire included 14 tasks with a context, all of which were presented in the text. In 26 cases, the reasons for liking or disliking the task were related to the context of the task. Most often this meant that the students liked the task because of the context. The written responses about the context tended to be short. Students wrote for example, “Burger,” “I like football,” and “because I like football.” Some more negative responses were “Mehh not a big fan of football” and “no football.”

The 14 context tasks involved contexts such as football, buying burgers, selling buns, and traffic jams. For Tasks 7 and 20, which are about football, the context of football was mentioned 10 times as a reason for liking the task, which is more often than it was the case for tasks with other contexts. The context of Task 4, which is about buying burgers, was mentioned second most often. It was brought up 4 times. The attractiveness of the football context was also shown in another way. Task 20 was given an appraisal score of 5 by 38.5% of the students which is the highest percentage of all the tasks in the study. For comparison, the average percentage for appraisal score 5 was 17.7%.

Although the specific context was overall a reason for liking the task, the text in which the task was presented was mostly a reason for disliking the task. Reactions were given such as “don’t like text tasks,” “too much text,” and “long text task.” Sometimes, but not often, text was described more positively. One student said, “I like the task because it is a text task.” Another student said, “a lot of text” while giving an appraisal score of 4. Sometimes it was clearly indicated that students liked some context tasks but did not like other context tasks because of the text. For example, one student made clear to like four tasks because they were about football, burgers, and discounts, but for four other context tasks, the student claimed to dislike them because of disliking text tasks.

**4.3.4 Reasons for liking or disliking different task types.** When comparing the reasons given for liking or disliking context tasks versus bare number tasks it was found that easiness was mentioned in 24.3% of the responses for bare number tasks and only in 13.7% of responses for context tasks. For the bare number tasks, this reason went together with a high appraisal score, and for the context tasks, the scores were a bit lower. A challenge as a reason was mentioned 3.7% for bare number tasks and 7.4% for context tasks. Mathematical content was mentioned 19.7% for the bare number tasks and 11.4% for the context tasks.

Differences were also found between the reasons for liking or disliking straightforward, puzzle-like, and estimation tasks. In the straightforward tasks, fun was only mentioned in 1.7% of the answers, while in the puzzle-like tasks and the estimation tasks this was 10.2% and 7.5% respectively. The reason given the most often for liking or not liking straightforward tasks was mathematical content. This was referred to in 21.7% of the responses. For the estimation tasks, the most common reason was challenge. This reason was mentioned in 15% of the tasks.

The estimation task elicited quite a number of reactions. In total, 13 students came up with a comment indicating that they were not familiar with estimation tasks. Five of them outrightly said that they did not know what estimation meant. These responses were mostly short, such as “I don’t know what estimate means.” Yet, one of these students came with a serious point of critique, saying, “there was a word that not everyone knows, it would have been easier to use other words like calculate instead of estimating it could be a bit difficult to figure out what that means for some people. And it also does not say how much a burger costs.” The remaining eight students seemed to indicate that they do not know what estimation means without outright stating this. In their responses, they pointed out that there was not enough information to solve the task by saying, “There was not enough information in the text, for example, how much a burger costs,” “I don’t have all the numbers.” One student made this lacking information very clear and commented, “How long is a car, the task could have given an average length of a car. All the information you need should be in the task in my opinion! I don’t know how long a car is ... 3m? 4m? 2m?.” This critique was also heard in the comment “All the information should be in the task, but maybe you’ll say that I know how much I sleep. No, maybe some nights 10 h. Other 8 h. And other nights 5 h. Then it will be difficult since I don’t write down how much I sleep each night to find the average. If you assume that you sleep 9 h each night, it is an ok task.”

For two students, this lacking information was not an issue. They were more positive about estimation tasks and seemed to know what estimation means. One student who gave an appraisal score of 5 said, “You can spend some time on it, and it is just you that can get to an answer that is correct for you.” Another student responded, “I like tasks where you have to estimate how much you do something like for example, how much electricity do you use, how many calories do you eat, and those kinds of things.” These two students seemed to enjoy exploration in mathematics and did not need there to be exact numbers in order to solve an estimation task.

#### **4.4 Students’ perceived solvability of mathematics tasks**

How difficult the students thought the tasks were, was measured by asking for each task whether they thought they would be able to solve the task. The perceived solvability scores ran from 1 to 5, where a score of 5 means that the student is very sure of being able to solve the task, and a score of 1 means that the student is very sure of not being able to solve the task.

As Table 6 shows the total mean score was close to 4, indicating that taking all cases together the students were rather sure that they were able to solve the tasks. However, the data also revealed that only in 45.1% of the cases, a score of 5 was given. For the scores of 4, 3, 2, and 1 this was 28.1%, 14.7%, 5.3%, and 6.9% respectively. The specific tasks used in this study are listed in Table 6b in the Appendix.



**Table 6.** Mean scores and standard deviations of the perceived solvability of the tasks.

	Task type	Mean score	SD
1	+ bare numbers <sup>a</sup>	4.72	0.775
2	- bare numbers	4.51	0.894
3	Fractions in context	3.99	1.037
4	Investigation with estimation in context	3.85	1.193
5	Puzzle-like problem bare numbers	4.45	1.01
6	+/ $\times$ in context	4.14	1.006
7	Puzzle-like problem in context	4.23	0.989
8	Fractions bare numbers	3.3	1.24
9	Percentages in context	3.61	1.239
10	Puzzle-like problem in context	4.08	1.168
11	Percentages bare numbers	4.03	1.145
12	Decimals in context	3.88	1.303
13	Puzzle-like problem in context	3.77	1.151
14	$\times$ bare numbers	4.28	1.111
15	Decimals in context	3.97	1.202
16	Puzzle-like problem bare numbers	4.15	1.265
17	Fractions bare numbers	3.74	1.207
18	Investigation with estimation context	3.82	1.286
19	Puzzle-like problem bare numbers	3.82	1.236
20	+/ $\times$ in context	4.33	1.085
21	: bare numbers	3.75	1.335
22	Fractions in context	3.59	1.231
23	Puzzle-like problem in context	3.83	1.294
24	Investigation with estimation in context	3.95	1.292
		3.99	1.154

The bare number addition task  $347 + 489$  (Task 1) got the highest perceived solvability score and the bare number fraction task the lowest. In general, both bare number tasks and context tasks were found in the tasks with a low perceived solvability score and in those with a high perceived solvability. Though with respect to the mathematical content, there was a division. Tasks about addition and multiplication were found rather solvable by the students, while tasks about fractions, percentages, and division were considered less or not solvable.

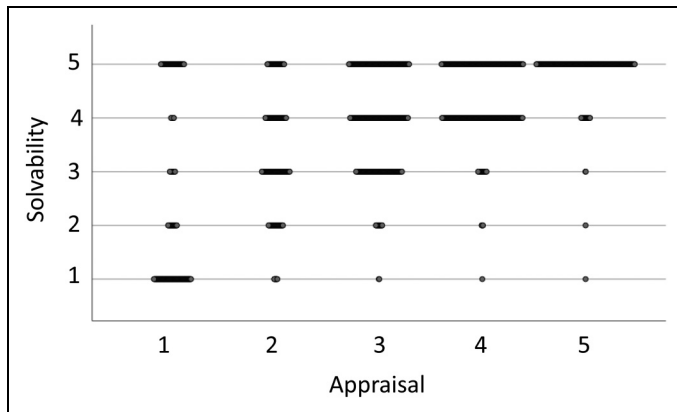
#### 4.5 Students' appraisal of mathematics tasks and the perceived solvability

In line with the earlier finding that the most common reason for liking or disliking a task was related to the easiness or difficulty of the task, it was no surprise that all the tasks that had the highest perceived solvability scores were also among the tasks with the highest appraisal scores (see Tables 3 and 6). Similarly, three of the five tasks with the lowest perceived solvability scores were also among the tasks that had the lowest appraisal scores.

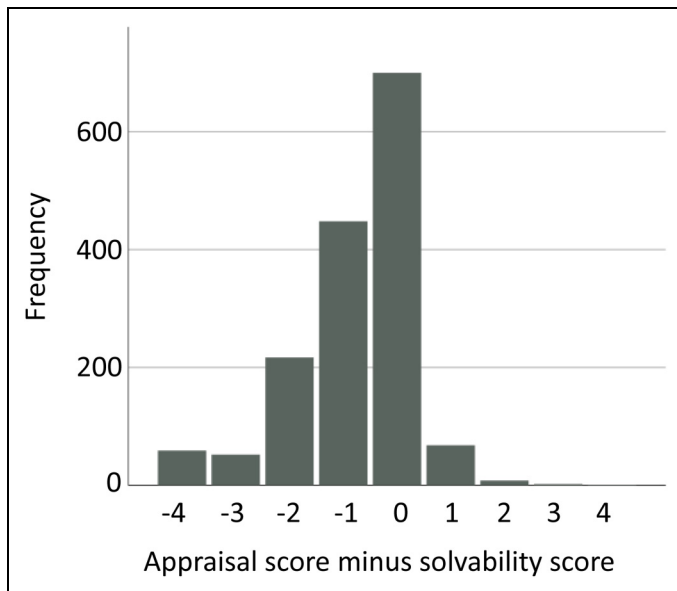
Further analysis of the data revealed that the student's appraisal score and the perceived solvability score were moderately correlated ( $\rho = 0.579, p < .01$ ), which is also shown in the contingency table (Table 7) and the scatter plot (Figure 4), which are based on all cases in which an appraisal score and a solvability score was given. The higher the students considered the solvability of the task the more they liked the task.

**Table 7.** Contingency table of the students' solvability and appraisal score for each task.

		Solvability					Total
		1	2	3	4	5	
Appraisal	1	94	23	14	8	59	198
	2	9	37	71	53	43	213
	3	2	17	116	146	152	433
	4	1	4	22	203	206	436
	5	1	1	2	23	248	275
Total		107	82	225	433	708	1555



**Figure 4.** Scatter plot of students' solvability score and appraisal score for each task.



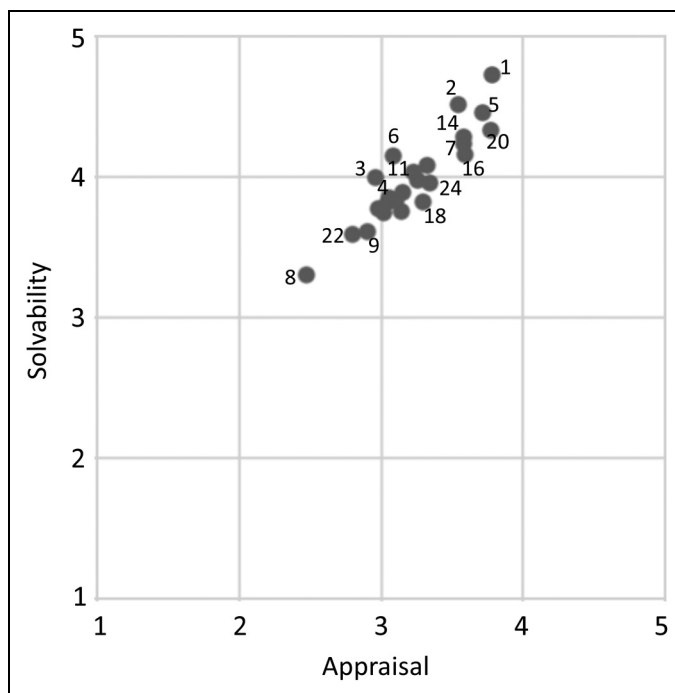
**Figure 5.** Frequency of the difference between the appraisal score and the solvability score.

This was also confirmed in another analysis in which the differences between these two scores were calculated. Figure 5 shows the frequencies of the differences between an appraisal score and a solvability score for all the cases in which the students gave both of these scores to a task. On the upper left side of the diagram (Figure 4), the solvability score is higher than the appraisal score and on the lower right side, the appraisal score is higher than the solvability score. The latter happened less frequently. In almost half of the cases, the score the students gave for liking/disliking the task had the same value as the score they gave for solvability.

When the solvability score was 1, the appraisal score was 1 in 87.9% of the cases, while in the case the solvability score was 5, the appraisal score was 5 in 35.0% of the cases. In the cases where the tasks were given an appraisal score of 1, the solvability score of 1 was given in 47.5% of those cases, while when the appraisal score was 5, the solvability score was 5 in 90.2% of the cases.

Finally, in order to know more about the tasks used in the questionnaire, we also investigated how the mean appraisal and solvability scores were related at task level. The results presented in the scatterplot of Figure 6 reveal that there was considerable variety in how the tasks were valued and that tasks which received about the same mean scores could be quite different in type and content.

For example, the group of seven tasks (Tasks 1, 2, 5, 20, 14, 7, and 16) with the highest mean appraisal and solvability score included three bare-number tasks but also three puzzle-like tasks. Three of these seven tasks were context tasks and four were bare number tasks. The tasks with the lowest mean appraisal and solvability score were two tasks about fractions (Tasks 8 and 22) and one about percentages (Task 9). Another noteworthy finding was that some apparently comparable tasks were differently judged. The mean solvability scores of two fractions in context tasks (Tasks 3 and 22) differed quite substantially due to the fractions ( $1/2$  and  $2/9$  respectively) involved in these tasks. A large difference was also found between the context (Task 9) and bare number task about percentages (Task 11) in the advantage of the bare number task. Furthermore, there were



**Figure 6.** Scatterplot of the mean appraisal and solvability score of each task.

differences between the two context tasks (Tasks 20 and 6) about addition and multiplication and between the tree estimation tasks (Tasks 4, 18, and 24).

## 5 Conclusion

### 5.1 Summary and reflection

In this study, we investigated students' opinions about mathematics tasks. The research questions focused specifically on what kinds of tasks students like or dislike, why they like or dislike the tasks, and whether this liking or not liking is related to whether students think they can solve the tasks. Answers to these questions were collected by means of an online questionnaire which was completed by 67 eighth-grade Norwegian students. The questionnaire consisted of 24 mathematics tasks varying in presentation, nature, and the operations and number types involved.

Taking together all the students' responses, then the overall picture the students had about the tasks was not that negative. The total mean appraisal score of the tasks came close to the score of 3, which means that the students on average did not really like the tasks very much, but they also did not dislike them very much. In addition to the tasks with an intermediate score, there were also tasks closer to the endpoints of the range. However, based on these tasks with a more outstanding appraisal, it was not possible to say which kinds of tasks were liked and which were not liked, because high appraisal scores were found for various sorts of tasks. These tasks included both bare number tasks with a straightforward problem as well as bare number tasks with a puzzle-like problem. Also, context tasks belonged to the highly appraised tasks. At the other end of the range, the results were clearer. The tasks that were rated lowest were particularly those about fractions and percentages. Considering the difficulties that many students have with fractions due to the whole-number bias (González-Forte et al., 2019; Reinhold et al., 2020; Sidney et al., 2021), it is not surprising that most of the tasks with the lowest appraisal scores had either fractions or percentages as a topic while all the highest appraised tasks were about whole numbers.

When comparing the appraisal scores for bare number tasks and context tasks we found that the students seemed to like the bare number task more, yet the difference was not statistically significant. An exception to this was when we compared the two presentation formats for the puzzle-like tasks. With respect to the reason for liking or disliking a task, easiness was more often mentioned for the bare number tasks than for the context tasks, while challenge as a reason was more frequently connected to context tasks.

Apart from this finding, students' opinions about context tasks revealed that the context can play a special role in whether students like or dislike a task. Often when the students said something about the context this was done in a positive way. A context for which this particularly was true was football, which is a very popular sport in Norway. In some towns with low populations, football is the only organized sport that children can participate in in their free time. Therefore, this is the sport that many students are the most familiar with, and many of them like it and play it after school. More in general, as it is emphasized by Clarke and Roche (2018), when contexts are about something the students are familiar with and in which they can use their own previous knowledge and experience, the students will be more engaged. The latter can contribute to students valuing the tasks as favorable, but it can also work in the opposite direction and can have a negative influence on students' performance (Almuna Salgado, 2017). It can also happen that students even indicate that they do not like the context of a task and that the teacher therefore needs to choose a different context. For example, such a situation was observed by Mack (1993). In her study, a student commented on a fraction problem that she does not like pizza and therefore never eats pizza, but instead likes ice cream. So, the teacher was asked to change the problem into an ice cream problem then it would be easier for the student. Also, several other authors have shown that the suitability of using

context in mathematics education is a complex issue, that if not used carefully, has the potential to restrict the mathematical development of some students (Sullivan et al., 2003). In line with this, Almuna Salgado (2017) even stated that the conflicting research outcomes about the use of context make clear that this issue is still not fully understood. An additional point that the students in our study raised in connection to the context was the use of text. Some students made it clear that the text was often a reason for disliking the task even when they liked the context.

When comparing the appraisal scores for the straightforward, the puzzle-like, and the estimation tasks we did not find a statistically significant difference, though the puzzle-like problems appeared to be more highly appraised by the students. Specifically, for the bare number tasks we found a significantly higher appraisal for the tasks with a puzzle-like problem than for the straightforward tasks. This preference for puzzle-like problems is in line with other studies that have shown that students were liking to be involved in challenging problem-solving tasks (Russo & Minas, 2020; Sullivan & Mornane, 2014) and that students were preferring open-ended context tasks to straightforward calculation tasks (Sullivan et al., 2013). Also, the reasons the students in our study gave for their appraisal scores suggested a positive attitude toward problem-solving tasks. For example, fun as a reason for liking a task was mostly mentioned in the case of the puzzle-like tasks, and for the estimation tasks the students often mentioned challenges. Another finding was that although in general, the reasons that were mainly mentioned for liking tasks were finding them to be easy and for disliking finding them to be difficult, the opposite was also the case. In some cases, the students made it clear that finding a task easy was the reason for not liking it that much. Likewise, it sometimes occurred that tasks that were found difficult or challenging were given a rather high appraisal score.

When it comes to the reasons given by the students for liking or not liking tasks, especially the estimation tasks elicited quite a number of comments varying from making clear that they did not know what estimation is, to criticizing how the tasks were formulated. According to the students who criticized the tasks, the necessary information for solving the tasks was lacking. In fact, these reactions revealed that the students were not familiar with estimation tasks. This however is not a surprise when we bring into mind what Sunde et al. (2022) found about this topic. As reported by them, Scandinavian schools, and particularly Norwegian schools, do not have much in the curriculum that is alluding to the topic of estimation. Since the students in our study were in the eighth grade when answering the questionnaire, it seems they have been through many years of schooling without having learned much about estimation. Therefore, the confusion that arose among the students around these tasks can aptly be regarded as a wake-up call for reconsidering what education has to offer students.

The last issue addressed in our study was whether the students thought they can solve the tasks and whether this was related to liking or disliking the tasks. The found mean score of 4 for solvability indicates that in general, the students were rather positive about whether they could solve the tasks. Yet, this score also means that the students were not very convinced of it. Only in about half of the cases, the students answered that they were very sure about being able to solve the tasks. As such, this finding could be considered a reasonable result, were it not that the students at the time of collecting the data were already in the eighth grade and that the tasks presented to them were mostly requiring only basic mathematical skills. This finding could again be a matter of what has been taught to the students, but there could also be other reasons for their reaction. It is very plausible that the students were quite able to solve the tasks but yet did not answer this because of low self-efficacy. As noted by Nilsen et al. (2022) the analysis of the TIMSS results has shown that the Norwegian students' self-concept in mathematics has declined from 2015 to 2019, which means that they feel less confident in this subject. This was also reflected in some utterances of the students where they called themselves "stupid" or otherwise seemed to feel helpless when doing mathematics.

When investigating whether the perceived solvability is related to the liking or disliking of tasks, it turned out that whether the students think they can solve a task might be a relevant factor in students' opinions about mathematics tasks. In our analyses, we found that students' perceived solvability of tasks and the appraisal scores given to the tasks were moderately positively correlated. The more the students expected that they could solve a task the more they liked the task. There seems to be a clear relationship, but at the same time, a further analysis carried out from the perspective of the tasks showed that there was considerable variety in how tasks were valued. Tasks with a particular appraisal and solvability score included quite different types of tasks, and similar types of tasks had different appraisal and solvability scores.

## 5.2 Limitations

Notwithstanding our research has led to interesting and inspiring, but challenging findings, the study certainly has some limitations. Therefore, the results should be viewed with caution. A limitation is that the students were only from four schools in the Nordland region of Norway and that only eighth-grade students were involved. Students in other schools and other grades may have different opinions about these tasks. From other studies, it is known that in lower grades students tend to report higher enjoyment and motivation in mathematics than in higher grades (Mazana et al., 2019; Russo & Minas, 2020).

Another issue that may have influenced our results is the used instrument. Although the online questionnaire was very feasible to reach the students and get the responses directly available for analyzing them, the use of this instrument also had its constraints, especially to get deeper knowledge about why students like or dislike tasks. The online answers were often very short and prompted the need for further questioning. Moreover, in about 40% of the cases, the students did not give a reason for their appraisal score. In this respect, interviewing the students would have been more appropriate in order to know more about how they think about the tasks. For example, learning more about why students think they can or cannot solve a task. Knowing this could provide more insight into how students approach specific tasks and what they consider as difficult and as easy in a task.

Also, the relation between the appraisal of the tasks and the perceived solvability of them needs more investigation. The responses of the students showed that these are not straightforwardly linked. For example, some students made it clear that they liked it when a task was difficult, but still answered that they thought they could solve it. Apparently, finding that a task is difficult does not automatically mean that the task cannot be solved. In connection with this, it would also be interesting to know more about the relation between the students' perceived solvability and the achieved solvability by also asking them to solve the tasks.

## 5.3 Further research

In accordance with the findings of Schukajlow (2015) and Sullivan et al. (2013), our study gives support to the view that investigations of various aspects of students' motivational orientation toward mathematics should not be limited to students' attitudes to mathematics in general. Besides knowing whether students like or dislike mathematics, there is another level of knowledge that is essential for making educational decisions for teaching, assessing, and designing mathematics education. Therefore, we think it is very important to include "mathematics tasks" as a factor in the theoretical models when investigating students' appraisal of mathematics. Moreover, besides working on more sensitive theoretical models to get a deeper understanding of the role of motivation in the learning of mathematics, knowledge about how students think about tasks is also of practical relevance. Like Middleton (2013), we think that it is the task where the motivation is triggered. Yet,

how many teachers, textbook authors, and assessment designers ever ask students which tasks they like? We do not know this, but our literature search showed that publications about students' opinions about mathematics tasks they come across in lessons, examinations, and textbooks, are very sparse. Following the current movement toward giving students a voice in mathematics education research (e.g., Hill & Seah, 2023; Røj-Lindberg & Hemmi, 2019; Roos, 2023; Zorn et al., 2022), we think this gap in literature needs to be filled by studies in which the perspective of students is taken. Earlier, Clarke et al. (2014, p. 119) made already a strong point for that by stating "research into the design and use of mathematical tasks in instructional settings must accommodate student intentions, actions and interpretations to at least the same extent as those of the teacher." In line with this statement, Johnson et al. (2017) even suggested adding a fifth area to Jones and Pepin's (2016) four areas which researchers should include in task design, namely students' experiences of tasks. This means that to enrich and deepen our understanding of teaching and learning mathematics we should give the students a voice in mathematics education research, particularly when this research is about the tasks that are used in mathematics education. Our article was meant to contribute to this.

Relevant in this respect is also that including students' feelings and thoughts about mathematics should go beyond whether they like/dislike mathematics in general. The latest TIMSS results (Mullis et al., 2020) have shown that half of the eighth graders in Norway do not like mathematics. However, for making instructional decisions it is important to know what this means at the task level. The TIMSS results also revealed that, apart from Singapore, the students in all Asian countries that participated in TIMSS disliked mathematics more than the international average. We therefore think that a study like we did would also be helpful to know more about Asian students' dislike of mathematics and to find leverage points for improvement.

## Acknowledgements

The authors thank the four teachers and their students involved in this study for their cooperation and contributions.

## Contributorship

Conceptualization: Ylva Høgset and Marja Van den Heuvel-Panhuizen; methodology: Ylva Høgset, Marja Van den Heuvel-Panhuizen, and Knut Berg; data collection: Ylva Høgset; resources: Ylva Høgset and Marja Van den Heuvel-Panhuizen; software: Ylva Høgset and Knut Berg; data curation: Ylva Høgset and Marja Van den Heuvel-Panhuizen; analysis: Ylva Høgset, Marja Van den Heuvel-Panhuizen, and Knut Berg; writing—original draft: Ylva Høgset; writing—review and editing: Marja Van den Heuvel-Panhuizen, Ylva Høgset, and Knut Berg; visualization: Marja Van den Heuvel-Panhuizen and Ylva Høgset; supervision: Marja Van den Heuvel-Panhuizen; project administration: Marja Van den Heuvel-Panhuizen and Ylva Høgset; funding acquisition: Ylva Høgset and Marja Van den Heuvel-Panhuizen.

## Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Nord Universitet (Grant No. 300027-153 for 2022).

## ORCID iD

Marja van den Heuvel-Panhuizen  <https://orcid.org/0000-0002-3985-7576>

## References

- Aguilar, J. J. (2021). High school students' reasons for disliking mathematics: The intersection between teacher's role and student's emotions, belief and self-efficacy. *International Electronic Journal of Mathematics Education, 16*(3), 1–11. <https://doi.org/10.29333/iejme/11294>
- Almuna Salgado, F. J. (2017). The role of context and context familiarity on mathematics problems. *Revista Latinoamericana de Investigación en Matemática Educativa, 20*(3), 265–292. <https://doi.org/10.12802/relime.17.2031>
- Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*(2), 117–148. [https://doi.org/10.1207/s15326985ep2802\\_3](https://doi.org/10.1207/s15326985ep2802_3)
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman and Co.
- Bernacki, M. L., & Walkington, C. (2018). The role of situational interest in personalized learning. *Journal of Educational Psychology, 110*(6), 864–881. <https://doi.org/10.1037/edu0000250>
- Clarke, D., & Roche, A. (2018). Using contextualized tasks to engage students in meaningful and worthwhile mathematics learning. *Journal of Mathematical Behavior, 51*(2), 95–108. <https://doi.org/10.1016/j.jmathb.2017.11.006>
- Clarke, D., Strømshag, H., Johnson, H. L., Bikner-Ahsbals, A., & Gardner, K. (2014). Mathematical tasks and the student. In P. Liljedahl, C. Nicol, S. Oesterle, & D. Allan (Eds.), *Proceedings of the 38th Conference of the International Group for Psychology of Mathematics Education and the 36th Conference of the North American Chapter of the Psychology of Mathematics Education* (Vol. 1, pp. 117–143). Vancouver, Canada: PME. <https://eric.ed.gov/?id=ED597799>
- Forsblom, L., Pekrun, R., Loderer, K., & Peixoto, F. (2022). Cognitive appraisals, achievement emotions, and students' math achievement: A longitudinal analysis. *Journal of Educational Psychology, 114*(2), 346–367. <https://doi.org/10.1037/edu0000671>
- Gojak, L. M. (2013). Motivation matters! *NCTM Summing Up*, 7 March. [https://www.nctm.org/News-and-Calendar/Messages-from-the-President/Archive/Linda-M\\_-Gojak/Motivation-Matters/](https://www.nctm.org/News-and-Calendar/Messages-from-the-President/Archive/Linda-M_-Gojak/Motivation-Matters/)
- González-Forte, J. M., Fernández, C., Van Hoof, J., & Van Dooren, W. (2019). Various ways to determine rational number size: An exploration across primary and secondary education. *European Journal of Psychology of Education, 35*(3), 549–565. <https://doi.org/10.1007/s10212-019-00440-w>
- Hannula, M. S. (2006). Motivation in mathematics: Goals reflected in emotions. *Educational Studies in Mathematics, 63*(2), 165–178. <https://www.jstor.org/stable/25472120> <https://doi.org/10.1007/s10649-005-9019-8>
- Hiebert, J., & Weame, D. (1993). Instructional tasks, classroom discourse and students' learning in second grade arithmetic. *American Educational Research Journal, 30*(2), 393–425. <https://doi.org/10.3102/00028312030002393>
- Hill, J. L., & Seah, W. T. (2023). Student values and wellbeing in mathematics education: Perspectives of Chinese primary students. *ZDM – Mathematics Education, 55*(2), 385–398. <https://doi.org/10.1007/s11858-022-01418-7>
- Jones, K., & Pepin, B. (2016). Research on mathematics teachers as partners in task design. *Journal of Mathematics Teacher Education, 19*(2), 105–121. <https://doi.org/10.1007/s10857-016-9345-z>
- Kunnskapsdepartementet. (2020). *Curriculum for mathematics year 1–10 (MAT01-05)*. Utdanningsdirektoratet. <https://www.udir.no/lk20/mat01-05?lang=eng>.
- Mack, N. (1993). Learning rational numbers with understanding: The case of informal knowledge. In T. Carpenter, E. Fennema, & T. Romberg (Eds.), *Rational numbers: An integration of research* (pp. 85–106). Lawrence Erlbaum Associates. <https://www.routledge.com/Rational-Numbers-An-Integration-of-Research/Carpenter-Fennema-Romberg/p/book/9781138984530>
- Mazana, M. Y., Montero, C. S., & Casmir, R. O. (2019). Investigating students' attitude towards learning mathematics. *International Electronic Journal of Mathematics Education, 14*(1), 207–231. <https://doi.org/10.29333/iejme/3997>
- McLeod, D. B. (1992). Research on affect in mathematics education: A reconceptualization. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics* (pp. 575–596). Macmillan.



- Middleton, J., Jansen, A., & Goldin, G. A. (2017). The complexities of mathematical engagement: Motivation, affect, and social interactions. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 667–699). National Council of Teachers of Mathematics.
- Middleton, J. A. (2013). More than motivation: The combined effects of critical motivational variables on middle school mathematics achievement. *Middle Grades Research Journal*, 8(1), 77–95. <https://www.infoagepub.com/mgrj-issue.html?i=p54c3b12e96247>
- Mitchell, M. (1993). Situational interest: Its multifaceted structure in the secondary school mathematics classroom. *Journal of Educational Psychology*, 85(3), 424–436. <https://doi.org/10.1037/0022-0663.85.3.424>
- Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D. L., & Fishbein, B. (2020). *TIMSS 2019 international results in mathematics and science*. Retrieved on 03.03.2023 from Boston College, TIMSS & PIRLS International Study Center. Retrieved March 3, 2023, from <https://timssandpirls.bc.edu/timss2019/international-results/>
- Nilsen, T., Kaarstein, H., & Lehre, A.-C. (2022). Trend analyses of TIMSS 2015 and 2019: School factors related to declining performance in mathematics. *Large-scale Assessments in Education*, 10(1), 15. <https://doi.org/10.1186/s40536-022-00134-8>
- Öztürk, M., Akkan, Y., & Kaplan, A. (2020). Reading comprehension, mathematics self-efficacy perception, and mathematics attitude as correlates of students' non-routine mathematics problem-solving skills in Turkey. *International Journal of Mathematical Education in Science and Technology*, 51(7), 1042–1058. <https://doi.org/10.1080/0020739X.2019.1648893>
- Reinhold, F., Obersteiner, A., Hoch, S., Hofer, S. I., & Reiss, K. (2020). The interplay between the natural number bias and fraction magnitude processing in low-achieving students. *Frontiers in Education*, 5(29), 1–13. <https://doi.org/10.3389/feduc.2020.00029>
- Roche, A., Gervasoni, A., & Kalogeropoulos, P. (2021). Factors that promote interest and engagement in learning mathematics for low-achieving primary students across three learning settings. *Mathematics Education Research Journal*, 35(3), 525–556. <https://doi.org/10.1007/s13394-021-00402-w>
- Røj-Lindberg, A.-S., & Hemmi, K. (2019). The students' perspective on school mathematics—A case study. In U. T. Jankvist, M. Van den Heuvel-Panhuizen, & M. Veldhuis (Eds.), *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education* (pp. 1962–1969). Utrecht: CERME. <https://hal.science/hal-02421574>
- Roos, H. (2023). Students' voices of inclusion in mathematics education. *Educational Studies in Mathematics*, 113(2), 229–249. <https://doi.org/10.1007/s10649-023-10213-4>
- Schukajlow, S. (2015). Effects of enjoyment and boredom on students' interest in mathematics and vice versa. In K. Beswick, T. Muir, & J. Wells (Eds.), *Proceedings of 39th Psychology of Mathematics Education Conference* (Vol. 4, pp. 137–144). Hobart, Australia: PME.
- Schukajlow, S., & Krug, A. (2014). Do multiple solutions matter? Prompting multiple solutions, interest, competence, and autonomy. *Journal for Research in Mathematics Education*, 45(4), 497–533. <https://doi.org/10.5951/jresmetheduc.45.4.0497>
- Schukajlow, S., Rakoczy, K., & Pekrun, R. (2023). Emotions and motivation in mathematics education: Where we are today and where we need to go. *ZDM – Mathematics Education*, 55(2), 249–267. <https://doi.org/10.1007/s11858-022-01463-2>
- Sidney, P. G., Thompson, C. A., Fitzsimmons, C., & Taber, J. M. (2021). Children's and adults' math attitudes are differentiated by number type. *The Journal of Experimental Education*, 89(1), 1–32. <https://doi.org/10.1080/00220973.2019.1653815>
- Skaalvik, E. M., Federici, R. A., & Klassen, R. M. (2015). Mathematics achievement and self-efficacy: Relations with motivation for mathematics. *International Journal of Educational Research*, 72, 129–136. <https://doi.org/10.1016/j.ijer.2015.06.008>
- Smith, M. S., & Stein, M. K. (1998). Selecting and creating mathematical tasks: From research to practice. *Mathematics Teaching in the Middle School*, 3(5), 344–350. <https://doi.org/10.5951/MTMS.3.5.0344>

- Stevens, T., Olivarez, A., Lan, W. Y., & Tallent-Runnels, M. K. (2004). Role of mathematics self-efficacy and motivation in mathematics performance across ethnicity. *Journal of Educational Research*, 97(4), 208–221. <https://doi.org/10.3200/JOER.97.4.208-222>
- Street, K. E. S., Stylianides, G. J., & Malmberg, L.-E. (2022). Differential relationships between mathematics self-efficacy and national test performance according to perceived task difficulty. *Assessment in Education: Principles, Policy & Practice*, 29(3), 288–309. <https://doi.org/10.1080/0969594X.2022.2095980>
- Sullivan, P., Clarke, D., & Clarke, B. (2013). Students' preferences for different types of mathematics tasks. In P. Sullivan, D. Clarke, & B. Clarke (Eds.), *Teaching with tasks for effective mathematics learning* (pp. 99–108). Springer. <https://link.springer.com/book/10.1007/978-1-4614-4681-1>
- Sullivan, P., & Mornane, A. (2014). Exploring teachers' use of, and students' reactions to, challenging mathematics tasks. *Mathematics Education Research Journal*, 26(2), 193–213. <https://doi.org/10.1007/s13394-013-0089-0>
- Sullivan, P., Zevenbergen, R., & Mousley, J. (2003). The contexts of mathematics tasks and the context of the classroom: Are we including all students?. *Mathematics Education Research Journal*, 15(2), 107–121. <https://doi.org/10.1007/BF03217373>
- Sunde, P. B., Petersson, J., Nosrati, M., Rosenqvist, E., & Andrews, P. (2022). Estimation in the mathematics curricula of Denmark, Norway and Sweden: Inadequate conceptualisations of an essential competence. *Scandinavian Journal of Educational Research*, 66(4), 626–641. <https://doi.org/10.1080/00313831.2021.1897881>
- Watson, A., & Mason, J. (2007). Taken-as-shared: A review of common assumptions about mathematical tasks in teacher education. *Journal of Mathematics Teacher Education*, 10(4), 205–215. <https://doi.org/10.1007/s10857-007-9059-3>
- Zorn, K., Larkin, K., & Grootenboer, P. (2022). Student perspectives of engagement in mathematics. In N. Fitzallen, C. Murphy, V. Hatisaru, & N. Maher (Eds.), *Proceedings of the 44th Annual Conference of the Mathematics Education Research Group of Australasia* (pp. 570–577). Adelaide, Australia: MERGA.

## Author biographies

**Ylva Høgset** finished her master's degree in mathematics education from Nord University in June 2022 and got a grant from the university to write this article based on her thesis. Currently, she is teaching seventh-, eighth-, and ninth-grade students at a Norwegian first–10th-grade school. A lot of the teaching she does is in special education mathematics.

**Marja van den Heuvel-Panhuizen** is a professor of mathematics education at Nord University, Norway, and is an emeritus professor of mathematics education at Utrecht University, the Netherlands. She has been working at the Freudenthal Institute since 1987, where she was appointed as a full professor in 2007. In 2015, she became emeritus and in 2017 she accepted a part-time position at Nord University. Her research interests lie with the instruction theory for mathematics education and the further development of the didactics of mathematics and its theoretical grounding. Her areas of research are assessment in mathematics education, primary school mathematics from K-G7, early algebra, higher-order thinking, digital media, textbooks, learning trajectories, mathematical potential of low achievers.

**Knut Berg** is a university lecturer at Nord University Business School. He is affiliated with the Division Economic Analysis and Accounting and the research group “Mathematics Teaching and Learning” of the Faculty of Education and Arts at Nord University. His research interests lie in mathematics education in upper secondary school and early university-level mathematics/statistics, and in algebraic topology.

## Appendix

**Table 3b.** Mean scores and standard deviation for the students' appraisal of tasks.

Task type	Task <sup>a</sup>	Mean score	SD	Absolute and relative frequency					
				Score:1	Score:2	Score:3	Score:4	Score:5	
1	+ bare numbers	347 + 489 =	3.7879	1.0886	4 6.1%	3 4.5%	14 21.2%	27 40.9%	18 27.3%
2	- bare numbers	912 - 677 =	3.5522	1.22206	6 9%	6 9%	17 25.4%	21 31.3%	17 25.4%
3	fractions in context	There are 12 bottles that has 1/2 litres each. How many litres does each person get if there are eight people sharing?	2.9701	1.1543	7 10.4%	18 26.9%	18 26.9%	18 26.9%	6 9%
4	investigation with estimation in context	Make an estimation of how many burgers you can buy for a million kroner? Explain how you came to your answer.	3.0625	1.31987	10 15.6%	11 17.2%	20 31.3%	11 17.2%	12 18.8%
5	puzzle-like problem bare numbers	Fill the missing digits in in the boxes $\square 3 - 2\square = 25$	3.7164	1.0236	6 9%	4 6%	12 17.9%	26 38.8%	19 28.4%
6	+/* in context	You are selling buns for 15kr, and adding jam costs an extra 5 kr. You bought ingredients for 132kr, and sold 84 buns. Every 4th bun had jam on it. How much money did you earn?	3.0909	1.19907	10 15.2%	8 12.1%	20 30.3%	22 33.3%	6 9.1%
7	puzzle-like problem in context	In a soccer tournament, teams get: 3 points for a win, 1 point for a tie, 0 points for a loss. Zedland has 11 points. What is the smallest number of games Zedland could have played?	3.5909	1.16325	5 7.6%	5 7.6%	18 27.3%	22 33.3%	16 24.2%
8	fractions bare numbers	$\frac{22}{15} \times \frac{45}{44} \times \frac{18}{10} =$	2.4848	1.09884	15 22.7%	18 27.3%	21 31.8%	10 15.2%	2 3%
9	percentages in context	The shop offers a 25% discount. The sale price of the binocular is 960 kr. What was the regular price?	2.9104	1.27602	13 19.4%	11 16.4%	19 28.4%	17 25.4%	7 10.4%

(continued)

Table 3b. (continued)

Task type	Task <sup>a</sup>	Mean score	SD	Absolute and relative frequency				
				Score:1	Score:2	Score:3	Score:4	Score:5
10 puzzle-like problem in context	A family has three children that have a total age of 36. The youngest child is half the age of the middle child, and the oldest child is three times the age of the youngest. How old are the children?	3.3333	1.37375	10 15.2%	8 12.1%	14 21.2%	18 27.3%	16 24.2%
11 percentages bare numbers	What is more? 15% of 750 or 35% of 350?	3.2273	1.28656	10 15.2%	8 12.1%	15 22.7%	23 34.8%	10 15.2%
12 decimals in context	The size of a carpet is 8.5 by 10 metres. 1.5 metres from the side is cut all around. What is the area in square metres of the reduced carpet?	3.1563	1.25	8 12.5%	11 17.2%	18 28.1%	17 26.6%	10 15.6%
13 puzzle-like problem in context	There is a row of books of different sizes on a shelf. There are 20 books to the left of the largest book and 22 books to the right of the smallest book. The largest book and the smallest book are both adjacent to the oldest one. What is the smallest possible number of books on the shelf?	2.9841	1.07	6 9.5%	13 20.6%	25 39.7%	14 22.2%	5 7.9%
14 × bare numbers	$14 \times 78 =$	3.5909	1.30062	6 9.1%	9 13.6%	11 16.7%	20 30.3%	20 30.3%
15 decimals in context	Rocco has 1.5 litres of orange soda and 2.25 litres of grape soda in his fridge. Antonio has 1.15 litres of orange soda and 0.62 litres of grape soda. How much more soda does Rocco have than Angelo?	3.2615	1.22827	8 12.3%	7 10.8%	21 32.3%	18 27.7%	11 16.9%
16 puzzle-like problem bare numbers	Fill the missing digits in the boxes. $243 + 1 \square 7 + \square 26 = 82$	3.6	1.18322	7 10.8%	3 4.6%	12 18.5%	30 40.2%	13 20%

(continued)

**Table 3b.** (continued)

Task type	Task <sup>a</sup>	Mean score	SD	Absolute and relative frequency				
				Score:1	Score:2	Score:3	Score:4	Score:5
17 fractions bare numbers	$\frac{13}{6} - \frac{7}{15} =$	3.0154	1.20536	9 13.8%	12 18.5%	20 30.8%	17 26.2%	7 10.8%
18 investigation with estimation in context	The traffic jam is 2 kilometres long. About how many cars do you think are in this traffic jam? Explain how you came to your answer.	3.3016	1.19993	7 11.1%	6 9.5%	22 34.9%	17 27%	11 17.5%
19 puzzle-like problem bare numbers	The difference between two double-digit numbers is 50. How many such pairs of double-digit numbers are there?	3.0455	1.20807	9 13.6%	11 16.7%	22 33.3%	16 24.2%	8 12.1%
20 +/× in context	A football team is participating in a cup. There are 14 players from this team participating. Registration fee for the team: 1700 kr. Accommodation and food per person: 900 kr. Renting a bus to and from the cup: 4600 kr. How much does it cost for the team to participate in the cup?	3.7846	1.25614	5 7.7%	5 7.7%	14 21.5%	16 24.6%	25 38.5%
21 ÷ bare numbers	$592 \div 37 =$	3.1538	1.26529	9 13.8%	9 13.8%	21 32.3%	15 23.1%	11 16.9%
22 fractions in context	$\frac{2}{9}$ of the people on a restaurant are adults. If there are 95 more children than adults, how many children are there in the restaurant?	2.8	1.09259	9 13.8%	14 21.5%	28 43.1%	9 13.8%	5 7.7%
23 puzzle-like problem in context	100 marbles are in row. Lisa starts on the left and takes 7 marbles each time, while Tim starts on the right and takes 3 marbles each time. They stop when there are no marbles left. How many marbles will each get?	3.1077	1.27626	11 16.9%	7 10.8%	20 30.8%	18 27.7%	9 13.8%

(continued)

**Table 3b.** (continued)

Task type	Task <sup>a</sup>	Mean score	SD	Absolute and relative frequency				
				Score:1	Score:2	Score:3	Score:4	Score:5
24	investigation with estimation in context Make an estimation of the hours you sleep in one year. Explain how you came to your answer. TOTAL	3.3492	1.32176	9 14.3%	6 9.5%	16 25.4%	18 28.6%	14 22.2%
		3.2455	1.25455	199 12.7%	213 13.6%	438 27.9%	440 28.1%	278 17.7%

<sup>a</sup>The tasks with the highest appraisal score are coloured blue and those with the lowest appraisal score are coloured orange.

**Table 6b.** Mean scores and standard deviations of the perceived solvability of the tasks.<sup>a</sup>

Task type	Task <sup>a</sup>	Mean score	SD
1	+ bare numbers $347 + 489 =$	4.72	0.775
2	- bare numbers $912 - 677 =$	4.51	0.894
3	fractions in context There are 12 bottles that has 1/2 litres each. How many litres does each person get if there are eight people sharing?	3.99	1.037
4	investigation with estimation in context Make an estimation of how many burgers you can buy for a million kroner? Explain how you came to your answer.	3.85	1.193
5	puzzle-like problem bare numbers Fill the missing digits in in the boxes $\square 3 - 2\square = 25$	4.45	1.01
6	+/ $\times$ in context You are selling buns for 15kr, and adding jam costs an extra 5 kr. You bought ingredients for 132kr, and sold 84 buns. Every 4th bun had jam on it. How much money did you earn?	4.14	1.006
7	puzzle-like problem in context In a soccer tournament, teams get: 3 points for a win, 1 point for a tie, 0 points for a loss. Zedland has 11 points. What is the smallest number of games Zedland could have played?	4.23	0.989
8	fractions bare numbers	3.3	1.24
9	percentages in context The shop offers a 25% discount. The sale price of the binocular is 960 kr. What was the regular price?	3.61	1.239
10	puzzle-like problem in context A family has three children that have a total age of 36. The youngest child is half the age of the middle child, and the oldest child is three times the age of the youngest. How old are the children?	4.08	1.168
11	percentages bare numbers What is more? 15% of 750 or 35% of 350?	4.03	1.145
12	decimals in context The size of a carpet is 8.5 by 10 metres. 1.5 metres from the side is cut all around. What is the area in square metres of the reduced carpet?	3.88	1.303

(continued)

**Table 6b.** (continued)

	Task type	Task <sup>a</sup>	Mean score	SD
13	puzzle-like problem in context	There is a row of books of different sizes on a shelf. There are 20 books to the left of the largest book and 22 books to the right of the smallest book. The largest book and the smallest book are both adjacent to the oldest one. What is the smallest possible number of books on the shelf?	3.77	1.151
14	× bare numbers	$14 \times 78 =$	4.28	1.111
15	decimals in context	Rocco has 1.5 litres of orange soda and 2.25 litres of grape soda in his fridge. Antonio has 1.15 litres of orange soda and 0.62 litres of grape soda. How much more soda does Rocco have than Angelo?	3.97	1.202
16	puzzle-like problem bare numbers	Fill the missing digits in in the boxes. $243 + 1\square7 + \square26 = 82$	4.15	1.265
17	fractions bare numbers	$\frac{13}{6} - \frac{7}{15} =$	3.74	1.207
18	investigation with estimation context	The traffic jam is 2 kilometres long. About how many cars do you think are in this traffic jam? Explain how you came to your answer.	3.82	1.286
19	puzzle-like problem bare numbers	The difference between two double-digit numbers is 50. How many such pairs of double-digit numbers are there?	3.82	1.236
20	+ / × in context	A football team of 14 players is participating in a cup. Registration fee for the team: 1700 kr. Accommodation and food per person: 900 kr. Renting a bus to and from the cup: 4600 kr. How much does it cost for the team to participate in the cup?	4.33	1.085
21	: bare numbers	$592 \div 37 =$	3.75	1.335
22	fractions in context	$\frac{2}{9}$ of the people on a restaurant are adults. If there are 95 more children than adults, how many children are there in the restaurant?	3.59	1.231
23	puzzle-like problem in context	100 marbles are in row. Lisa starts on the left and takes 7 marbles each time, while Tim starts on the right and takes 3 marbles each time. They stop when there are no marbles left. How many marbles will each get?	3.83	1.294
24	investigation with estimation in context	Make an estimation of the hours you sleep in one year. Explain how you came to your answer.	3.95	1.292
		TOTAL	3.99	1.154

<sup>a</sup>The tasks with the highest perceived solvability are coloured blue and those with the lowest perceived solvability are coloured orange.