Adaptive number knowledge among primary school students of various ages

SALLA PEHKONEN, ANTTI LEHTINEN, PASI NIEMINEN AND MARKUS HÄHKIÖNIEMI

Previous studies have highlighted the importance of primary school students' adaptive number knowledge, which includes knowledge of numerical characteristics and relations. In this study, students from the second (aged 8), fourth (aged 10), and sixth (aged 12) grades ($n = 205$) answered a modified version of the *Arithmetic production task*, which has been used to measure primary school students' adaptive number knowledge. Significant differences among grade levels were found. In addition, a latent profile analysis revealed four profiles based on the students' answers, with profile membership being associated with grade level. Similar to previous research, the current study found evidence of both age-dependent and individual differences. Furthermore, an analysis of the strategies used by the students to produce solutions revealed differences among the profiles.

Fluency in arithmetic is one of various components of the arithmetic skills that students need to acquire. Apart from mastering calculations, arithmetic skills also include the ability to flexibly utilize multiple strategies and adapt one's knowledge to a range of mathematical problems. Distinguishing and using relations between numbers and operations in various ways is an essential element of adaptivity in arithmetic, which McMullen et al. (2016) called adaptive number knowledge. In order to measure adaptive number knowledge, McMullen et al. (2016) created a test where students are asked to produce mathematical expressions, that is, they have to recognize and use the relations between numbers and operations in multiple ways in order to reach their solutions. This measure provides information about individual differences, and the results are associated with pre-algebra skills (McMullen et al., 2017). Thus

Salla Pehkonen, *University of Jyväskylä*
Antti Lehtinen, *University of Jyväskylä*
Pasi Nieminen, *University of Jyväskylä*
Markus Hähkiöniemi, *University of Jyväskylä*

Pehkonen, S., Lehtinen, A., Nieminen, P. & Hähkiöniemi, M. (2022). Adaptive number knowledge among primary school students of various ages. *Nordic Studies in Mathematics Education*, 27 (4), 3–24.

far, however, the focus has been more on older primary school students and the operations they use. More information is needed on how adaptive number knowledge develops with age. In addition, when students produce multiple solutions, one of them may be used as a template to create others, for example, by replacing addition with multiplication. Recognizing these kinds of strategies as a way of modifying solutions that are yet be studied may help us understand differences between students. The current study aims to narrow these research gaps.

Theoretical background

Flexible mathematical thinking

Flexible thinking plays a significant role in the application of a range of strategies in solving mathematical problems. Many studies have defined flexible mathematical thinking as competence in the use of various strategies and being able to choose the most suitable one for the given situation (Blöte et al., 2000; Hästö & Palkki, 2019; Star & Newton, 2009; Threlfall, 2002; Verschaffel et al., 2009). In the literature, the terms flexibility and adaptivity are sometimes used synonymously, yet they tend to convey different meanings. Although the definition of flexible mathematical skills includes the ability to choose the most appropriate strategy, many studies (e.g. Heinze et al., 2009; Verschaffel et al., 2009) have defined flexibility as the ability to use multiple strategies and adaptivity as the ability to choose the most appropriate strategy. Hatano (2003) defined adaptivity as inclusive of flexibility and creativity, and as Selter (2009) described: "Adaptivity is the ability to creatively develop or to flexibly select and use an appropriate solution strategy in a (un)conscious way on a given mathematical item or problem, for a given individual, in a given sociocultural context" (p. 624). Flexible thinking and adaptive use of various strategies have laid a strong foundation for mathematical skills due to the cumulative nature of mathematics, and previous knowledge has often been used in multiple ways and modified to answer new kinds of problems.

Flexible mathematical thinking is related to the efficiency involved in obtaining solutions, making strategy-based choices, and the understanding behind these actions. Generally, speed and accuracy are common indicators of mathematical skills because they depict one's efficiency. Hästö and Palkki (2019) determined whether students' tendency to use standard or innovative strategies is related to their speed and accuracy. As a standard strategy, they referred to the mechanical strategies illustrated by the teacher. Conversely, innovative strategies do not follow specific steps and are adaptively formed from number relations and characteristics.

According to their results, students' competence in using innovative strategies correlates with their speed and accuracy. Thus, choosing a suitable strategy can make a difference in response speed and accuracy.

If a student knows multiple strategies to similar problems, they have to choose which strategy to use every time a similar problem is encountered. When the student has learned to use a certain strategy to solve certain types of problems, their speed and accuracy evolve even though the strategy does not (Lemaire & Siegler, 1995). However, learning to use only one strategy for one problem type may lead to unconnected knowledge, that is, knowledge that cannot be flexibly deployed for a larger variety of problems (Gravemeijer et al., 2016). Furthermore, a student's age is related to their strategy usage; when solving missing-value problems, younger students tend to use more additive strategies, but as they grow older, their use of additive thinking decreases, and they begin to deploy multiplicative strategies (Van Dooren et al., 2010).

Being flexible in mathematics requires conceptual and procedural knowledge of numbers and operations as well as the competence to shift among multiple strategies (Siegler & Lemaire, 1997). The ability to use a certain procedure or strategy is not the only factor that relates to one's mathematical skills; being able to understand the underlying idea behind the procedure or strategy is also fundamental. Hatano (2003) pointed to two kinds of expertise: routine and adaptive. Routine expertise is memorized knowledge wherein one can mechanically use familiar strategies in familiar situations. Adaptive expertise is knowledge of procedures and concepts and the ability to recognize the relations between them, thereby enabling the use of novel strategies in new situations. When solving arithmetic problems, adaptive expertise manifests as the adaptive and flexible use of strategies in number calculations (Brezovszky et al., 2015).

Adaptive number knowledge

Students' adaptive use of numbers can be measured through adaptive number knowledge, which focuses on students' competence in handling a range of number characteristics. It is defined as the ability to distinguish and use the relations between numbers and operations in the most advantageous way for the given situation (Brezovszky et al., 2015; McMullen et al., 2016). According to McMullen et al. (2016, 2017, 2019), adaptive number knowledge forms the foundation for the development of one's mathematical thinking and is a valid descriptor of one's mathematical competence since it is associated with arithmetical fluency and pre-algebra skills.

Adaptive number knowledge can be measured with the *Arithmetic production task* (APT) (Brezovszky et al., 2015, 2019; McMullen et al., 2016, 2017, 2019). The premise behind the test is to produce as many solutions of a given value as one can in the given time by using only the given numbers. The items are divided into dense and sparse categories according to the arithmetic relations between the numbers in them. In the dense items, the given and target number are selected so that there are many common factors and multiples, and therefore, more straightforward solutions can be produced. In the sparse items, however, there are only a few or no common factors or multiples. McMullen and colleagues (2017, 2019) analyzed students' solutions by differentiating between simple solutions containing only additive operations (e.g. $3 + 4 - 1$), or only multiplicative operations (e.g. $2 \cdot 3$), or complex solutions containing both (e.g. $2 \cdot 4 - 2$). Thus, the students' solutions could be categorized as follows: simple solutions in the dense items, simple solutions in the sparse items, complex solutions in the dense items, and complex solutions in the sparse items. Since complex solutions require connecting operations with different characteristics, they are considered more adaptive.

A few studies have examined and compared adaptive number knowledge of primary school students of various ages (McMullen et al., 2016, 2017). These studies employed the same methods and definitions but slightly different analyses, and there was some variation between the results. McMullen et al. (2017) divided the solutions of fourth to sixth graders into simple and complex solutions and studied these solutions in the dense and sparse items. They found significant grade-level differences in the simple solutions in the dense items and in the complex solutions in both item types. A study by McMullen et al. (2016) showed no significant differences among third to fifth graders when the variables in the analysis were the total number of correct solutions, the total number of complex solutions, and the proportion of complex solutions per correct solutions. The varying results regarding grade-level differences raise the question of whether adaptive number knowledge develops throughout primary school and can be detected in students who are novices at arithmetic operations.

While the number of simple and complex solutions performed relates to students' adaptive number knowledge, the focus has hitherto been on how students recognize and use the relations between the given numbers and connect them with basic operations to produce the target number. There has been some discussion about the different solution strategies that students employ (McMullen et al., 2017), but observations have only focused on the number of simple and complex solutions. In this study, we focus on the strategies that students use to modify their solutions, that is, how they recognize relations between different solutions and use them to produce more solutions. For example, after a student has produced a solution $6+6-1$, they may relate $6+6$ to $2\cdot 6$ and produce $2\cdot 6-1$. This way, they produce two related solutions. Related solutions in the APT have not been studied but they add another layer of relations that play a role in the APT and partly explain how students can create solutions.

Research questions

To study students' adaptive number knowledge, the age range was expanded to include differences among second, fourth, and sixth graders. Since the test has not previously been used on second graders, who are just learning multiplication, we had to modify it. The modified version of the APT (Brezovszky et al., 2015, 2019; McMullen et al., 2016, 2017, 2019) is called the *Small number arithmetic production task* (s-APT). Our modification raises the question of whether the modified test can detect individual differences as in the original test (McMullen et al., 2017). In addition, we conducted a new analysis for the students' solutions in order to study how these students produced solutions by utilizing the relations between them. The research questions are as follows.

1. What differences exist among second, fourth, and sixth grade students in their simple and complex solutions in the dense and sparse items?
2. What latent profiles describe second, fourth, and sixth grade students' adaptive number knowledge?
3. What are the related solutions in the students' answers, and how are these related to the differences between the latent profiles?

Methods

Participants

The participants were 205 students from the second ($n=74$, girls = 35), fourth ($n=59$, girls = 26), and sixth ($n=72$, girls = 38) grades. The mean ages of the second, fourth, and sixth graders were eight years and five months ($SD=4.0$ months), 10 years and four months ($SD=3.6$ months), and 12 years and five months ($SD=3.6$ months), respectively. The participating classes were from three suburban schools in Finland. Four classes from each grade participated, so there were 12 classes in total. We made sure that all the classes were at least familiar with multiplication since the order of contents in the lessons varied depending on the materials used by the teachers. The teachers volunteered their classes for the study, and informed consent forms were obtained from the students' legal guardians. All students participated in the test, but for those who did not want to be involved in the study, the results relating to them were not analyzed, and their data were destroyed.

Procedures

The data were collected at the end of the fall term of 2019. The test was performed in a paper–pencil format in a classroom environment administered by a researcher. Before starting the test, the administrator explained the assignment without any prompts about what kinds of solutions were expected, and the students got to practice with a sample item. Before working with the actual research items, the students were allowed to ask questions. For each item, the administrator turned on a timed presentation that gave a voice signal when the time was up. The test administrator was the same for all classes; thus, the assignment was equal for everyone.

Measures

The students' adaptive number knowledge was measured with the s-APT. The same test was used for all the participants in order to obtain comparable data. The numbers in the items had to be small enough, and there was a chance in every item for the second graders to produce multiplications (given that the second graders' curriculum only contained multiplication with the numbers one to five and 10 (Opetushallitus, 2016)).

The test contained four items, each item with a 90-second time limit. Each item included four or five given numbers that the students were supposed to use in their solutions as many times as they wanted and a target number that had to be the value of every solution. The permissible operations were addition, subtraction, multiplication, and division, which could be used as many times as the students wanted. The goal was to create as many solutions as possible within the 90 seconds. Cronbach's alpha reliability for the total number of correct solutions across the four items was .86.

The items can be categorized into two types according to the number of multiples and common factors that can be found among the given numbers and the target number. Given the higher number of these characteristics between the numbers in the dense items, it is possible to find many straightforward solutions. In the sparse items, the number of these characteristics is significantly lower. Thus, in the sparse items, only three or two one-step solutions can be found. Meanwhile, in the dense items, five or seven one-step solutions can be found (see table 1). However, there are infinitely many solutions when using more than one step.

Table 1. *Given numbers, target numbers, item types, and number of one-step solutions in the APT and s-APT*

Item	Given numbers in the APT	Target numbers in the APT	Given numbers in the s-APT	Target numbers in the s-APT	Type	Number of 1-step solutions in the APT	Number of 1-step solutions in the s-APT
example	1, 2, 3, 4	6	1, 2, 3, 4	6	dense	5	5
1	2, 4, 8, 12, 32	16	2, 4, 8, 12, 32	16	dense	7	7
2	1, 2, 3, 5, 30	59	1, 2, 6, 14, 42	8	sparse	0	3
3	2, 4, 6, 16, 24	12	2, 4, 6, 16, 24	12	dense	5	5
4	3, 5, 30, 120, 180	12	2, 3, 4, 17	5	sparse	0	2

The dense items used in this study were the same as those used in McMullen et al. (2017, 2019), but the sparse items had to be modified into more suitable items for the second graders. One item that could be used as a sparse item for the second graders was taken from Brezovszky et al. (2019). The second sparse item was composed of small numbers and a few one-step solutions. Table 1 shows both the given and target numbers from the original version of the APT (Brezovszky et al., 2019; McMullen et al., 2017) and the given and target numbers of the modified s-APT. The item types and the number of one-step solutions in the APT and s-APT are also shown.

Figure 1 illustrates all the one- and two-step solutions in a sparse item in the APT and a sparse item in the s-APT using the same representation as in McMullen et al. (2016). It shows that the sparse item in the s-APT was a bit easier since there were two one-step solutions and some two-step solutions with only additive operations; the other two-step solutions included both additive and multiplicative operations. However, there were no one-step solutions in the sparse item of the APT, and the two-step solutions included only multiplicative operations or both additive and multiplicative operations.

Analysis

In the test results, only the correct solutions were analyzed. Correct solutions were defined as expressions wherein the given numbers were used and the value was the target number. Therefore, commutative solutions (e.g. $2 + 4$ and $4 + 2$) were accepted as different solutions. In addition, the expressions where brackets had to be used (for example $2 + 4 \cdot 2 = 12$) were considered as correct solutions because brackets are not taught in the

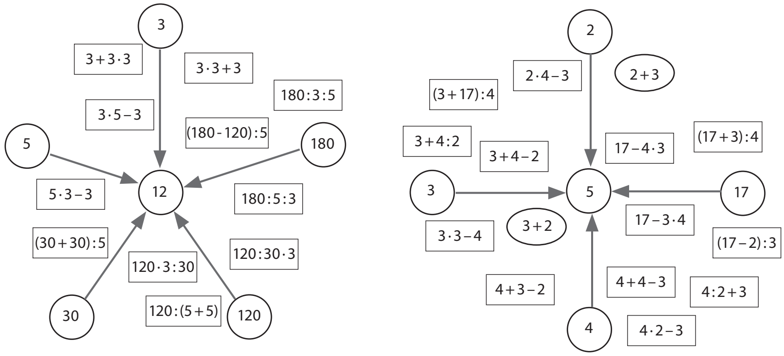

Figure 1. *A sparse item from the APT (left) and a sparse item from the s-APT (right) are presented as numerical networks*

Note. The number in the middle is the target number, which is surrounded by the given numbers. The one-step solutions are circled, while the two-step solutions are in rectangles.

second grade. The correct solutions were coded as simple solutions or complex solutions; simple solutions contain only additive or only multiplicative operations, while complex solutions contain both (McMullen et al., 2017, 2019).

The number of simple and complex solutions was counted from the students' answers. The simple and complex solutions were first compared between grade levels on an item-by-item basis, by using a one-way ANOVA, to confirm that no items were either too easy or difficult for the students. Thereafter, one-way ANOVAs were used to compare the simple and complex solutions between the grade levels in the dense and sparse items separately. A latent profile analysis (LPA), which is the same as latent class analysis (LCA), except that the variables are continuous instead of binary, was conducted to understand the way in which the students were divided into profiles based on their test performance. The LPA divides a heterogeneous group of participants into homogeneous groups based on the continuous variables (Berlin et al., 2014; Muthén & Muthén, 2017). Several statistical indicators were observed to estimate the models: the Aikake information criterion (AIC), the Bayesian information criterion (BIC), entropy, and a p-value for the bootstrapped likelihood ratio test (BLRT p-value). The students were assigned to profiles based on the most likely class membership. The variables in the LPA were simple solutions in the dense items, simple solutions in the sparse items, complex solutions in the dense items, and complex solutions in the sparse items.

After dividing the students into profiles, we compared their answers and searched for related solutions that might explain their strategies to modify solutions. By related solutions, we mean a pair of solutions where the second solution could be produced by chancing a term or terms from the first solution. Otherwise, these solutions must be identical. If a relation requires two or more terms, these terms must be situated next to each other so that the change is noticeable. For instance, when a student uses decomposition, the decomposed numbers should be next to each other ($12+\mathbf{4}$ and $12+\mathbf{2}+\mathbf{2}$ is acceptable but $12+\mathbf{4}$ and $\mathbf{2}+12+\mathbf{2}$ is not). It matters not the order in which these solutions are presented on the paper or whether there are other solutions between the pair. In addition, one solution can be related to different solutions by different relations. For example, the solution $2\cdot 6$ is related to $6\cdot 2$ by commutative law, but $2\cdot 6$ is also related to $6+6$ by repeated addition.

The first author analyzed all the solutions and defined the relations found. Another researcher coded 10% of the data and sought to identify new possible relations. Their codes were compared, which led to an adjustment of some definitions and the addition of a new relation. Thereafter, another round of coding was conducted, and inter-rater reliability was evaluated with Krippendorff's alpha ($\alpha=.91$), which suggested high agreement among the raters since the value exceeded .80 (Krippendorff, 2004). Next, a one-way ANOVA was conducted to compare the related solutions among the profiles.

Results

Item-by-item differences between grade levels

The means and standard deviations for the simple and complex solutions for each item for the second, fourth, and sixth graders are presented in table 2. For the simple solutions, one-way ANOVA (see table 2) indicated statistically significant differences between grade levels in the first ($p<.001$), second ($p=.002$), and third ($p<.001$) items. However, there were no significant differences in the fourth item ($p=.463$). The Tukey post-hoc test results are presented via indexes in table 2, and they indicate significant differences between the fourth and sixth graders only in the first item. In summary, in the first item, there were significant differences among all grade levels, while in the second and third items, the second graders had significantly fewer simple solutions than the other students. In the fourth item, no significant differences were found. For the complex solutions, the results from the one-way ANOVA (see table 2) indicate statistically significant differences in all the items (I1: $p<.001$; I2: $p<.001$; I3: $p<.001$; and I4: $p<.001$). However, according to the Tukey post-hoc test (see table 2), the sixth graders outperformed the second and fourth graders in the first three items, and there were significant differences among all grades in the fourth item.

Table 2. *Descriptive statistics and values from one-way ANOVAs and post-hoc tests of simple and complex solutions for Items 1, 2, 3 and 4*

	2nd Grade M (SD)	4th Grade M (SD)	6th Grade M (SD)	Total M (SD)	F(2, 202)	*p*	η^2
I1 Simple	2.12[a] (1.20)	3.29[b] (1.59)	4.49[c] (1.77)	3.29 (1.82)	43.48	< .001	.30
I2 Simple	2.31[a] (.98)	2.90[b] (1.14)	3.01[b] (1.58)	2.73 (1.30)	6.42	.002	.06
I3 Simple	2.70[a] (1.51)	4.02[b] (1.44)	4.40[b] (1.78)	3.68 (1.75)	22.69	< .001	.19
I4 Simple	1.91[a] (1.04)	2.02[a] (1.12)	2.15[a] (1.41)	2.02 (1.20)	.77	.463	.01
I1 Complex	.00[a] (.00)	.20[a] (.45)	.46[b] (.75)	.22 (.54)	15.10	< .001	.13
I2 Complex	.04[a] (.26)	.22[a] (.49)	.74[b] (1.11)	.34 (.79)	17.72	< .001	.14
I3 Complex	.03[a] (.16)	.20[a] (.45)	.69[b] (.87)	.31 (.64)	26.07	< .001	.21
I4 Complex	.03[a] (.16)	.47[b] (.73)	.89[c] (1.12)	.46 (.85)	22.50	< .001	.18

Note. There are statistical differences only between the profiles that do not have the same letter in the index

Table 3. *Descriptive statistics and values from one-way ANOVAs and post-hoc tests of simple solutions in the dense and sparse items and complex solutions in the dense and sparse items by grade level*

	2nd Grade M (SD)	4th Grade M (SD)	6th Grade M (SD)	Total M (SD)	F(2, 202)	*p*	η^2
Dense Simple	4.82[a] (2.39)	7.31[b] (2.55)	8.89[c] (3.04)	6.97 (3.18)	42.65	< .001	.30
Sparse Simple	4.22[a] (1.75)	4.92[a, b] (1.92)	5.17[b] (2.57)	4.75 (2.15)	3.91	.022	.04
Dense Complex	.03[a] (.16)	.41[b] (.70)	1.15[c] (1.40)	.53 (1.03)	28.33	< .001	.22
Sparse Complex	.07[a] (.38)	.69[b] (1.04)	1.63[c] (2.05)	.80 (1.50)	24.39	< .001	.19

Note. There are statistical differences only between the profiles that do not have the same letter in the index

Differences in the dense and sparse items between grade levels

When the dense (I1 & I3) and sparse (I2 & I4) items were taken into account, there were four variables: simple solutions in the dense items (DS), simple solutions in the sparse items (SS), complex solutions in the dense items (DC), and complex solutions in the sparse items (SC). Table 3 presents the means and standard deviations and the statistical indicators from the one-way ANOVAs for these variables. There were statistically significant differences among the grade levels in all variables (DS: $p<.001$; SS: $p=.022$; DC: $p<.001$; and SC: $p<.001$). A multiple comparison test was run, and the Tukey post-hoc test results are presented via indexes in table 3. They indicate that the second graders had significantly lower means than the sixth graders in every variable. However, the fourth graders' simple solutions in the sparse items did not differ significantly from those of either the second or sixth graders'. In other variables (i.e. DS, DC, and SC), there were significant differences among all grades. We also checked whether the students' gender or class had an influence on their performance, but no notable effects were found.

Student profiles

Previous studies regarding students' adaptive number knowledge (McMullen et al., 2017, 2019) have included a narrower age range and older students. Thus, we did not use existing study profiles to conduct a confirmatory LPA. Instead, an LPA was run to determine the kinds of profiles formed when second graders are also included. After the first round of the LPA, one sixth grade student had to be excluded because of superior performance on the test. The student had their own profile wherein no other student was included, and therefore, the LPA was run again without them.

When observing entropy values, the higher the value is on a scale from 0 to 1, the better. When the value exceeds .80, the classes are sufficiently discriminatory (Tein et al., 2013). The BIC is considered a better indicator than other IC indicators (Nylund et al., 2007), and the value should be as small as possible. The BLRT yields a p-value, that is, the result is significant if the value is $< .05$. It is not frequently used as an indicator, but it has been shown to be the best likelihood-based indicator even more precise than the BIC (Nylund et al., 2007). Therefore, the number of classes wherein the BIC is as low as possible, entropy is as high as possible, and BLRT p-values under .05 are considered best.

The statistical indicators from the LPA are presented in table 4, with the chosen four-class model bolded. Although the BIC value was the lowest and the BLRT p-value was significant in the seven-class model, the entropy value did not exceed .80. According to Nylund et al. (2007), when the BLRT p-value is nonsignificant, it remains nonsignificant after increasing the number of classes. According to table 4, this does not seem

Table 4. *Statistical Indicators from the LPA*

Number of classes	AIC	BIC	Entropy	BLRT
2	2928.87	2972.01	.95	.01
3	2875.16	2934.89	.92	.01
4	2848.20	2924.52	.80	.01
5	2852.55	2945.46	.70	.60
6	2862.47	2971.97	.58	.99
7	2744.39	2870.48	.73	.01
8	2754.00	2896.68	.65	.97

to be the case, but Nylund et al. also pointed out that with a likelihood-based test, one should stop increasing the number of classes after the first nonsignificant *p*-value. The three-class model would also have been a good choice, but in this case, over two-thirds of the students would have been in the same profile.

The four-class model contains four profiles: Basic, Average, High-Simple, and High-Complex. The mean values of each variable for all the profiles are presented in figure 2. For the students in the Basic profile, the mean values of each variable were lower than the average. In addition, the number of complex solutions was very low in both item types. For the students in the Average profile (which would have been integrated into the Basic profile in the three-class model), the number of simple solutions was somewhat above average, but the number of complex solutions is somewhat below average in both the dense and sparse items. In addition, the Average profile was the only profile wherein the students indicated more complex solutions in the dense items than in the sparse items. For the students in the High-Simple profile, the number of simple solutions was substantially higher than the average (higher than in all other profiles), and the number of complex solutions was around average in both the dense and sparse items. Finally, for the students in the High-Complex profile, the number of complex solutions was considerably high. In addition, the number of solutions in all the variables was above average, but the number of simple solutions in the sparse items was still lower than in the Average and High-Simple profiles.

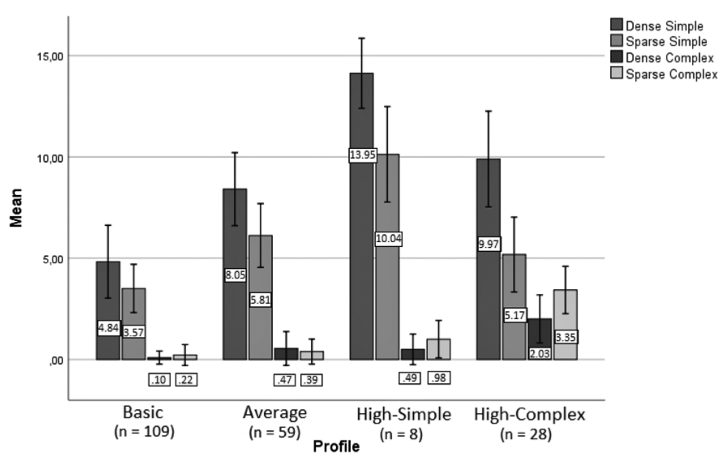

Figure 2. *Student profiles and the mean values and standard deviations for each variable in the profiles*
Note. Error bars = ± 1 SD

Table 5. *Number of students and percentage share per profile for each grade*

	Profiles				
Grades	Basic (%)	Average (%)	High-Simple (%)	High-Complex (%)	Total (%)
2	59 (79.7)	13 (17.5)	1 (1.4)	1 (1.4)	74 (100.0)
4	28 (47.5)	24 (40.7)	2 (3.4)	5 (8.4)	59 (100.0)
6	22 (31.0)	22 (31.0)	5 (7.0)	22 (31.0)	71 (100.0)
Total	109 (53.5)	59 (28.9)	8 (3.9)	28 (13.7)	204 (100.0)

The number of students per profile is presented in table 5. Most second and fourth graders belonged to the Basic (79.7 % and 47.5 %, respectively) and Average (17.5 % and 40.7 %, respectively) profiles. The sixth graders were spread evenly into the Basic, Average, and High-Complex profiles (31.0 % in each). The High-Simple profile was very small and included only 3.9 % of the participants. Due to the small number of students in the High-Simple profile, a chi-squared test could not be conducted, but Fisher's exact test indicated a significant positive association ($p < .001$) between grade level and profile membership.

Related solutions

We found eight relations from the students' answers: decomposition, using a difference, commutative law, repeated addition, adding a zero, multiplying by one, factorization and using a quotient. The definitions and examples of these relations are presented in table 6. The average number of relations in the test was 1.19 ($SD = .97$) for the students in the Basic profile. The mean values for the Average, High-Simple, and High-Complex profiles were 2.64 ($SD = 1.01$), 3.50 ($SD = .76$), and 3.71

Table 6. *Definitions and examples of the relations*

Relation	Definition	Example
Decomposition	A number is decomposed into smaller numbers that are then added up. It can also be used in subtraction.	$6+2$, $2+4+2$ and $14-6$, $14-2-2-2$
Using a difference	A number is presented as the difference of other numbers.	$2+3$ and $4-2+3$
Commutative law	The order of the operands is changed, but the terms stay the same.	$6 \cdot 2$ and $2 \cdot 6$
Repeated addition	A multiplication is presented as a repeated addition	$4 \cdot 4$ and $4+4+4+4$
Adding a zero	The same number has been added and then subtracted or vice versa.	$2+3$ and $17-17+2+3$
Multiplying by one	A solution includes a multiplication by one. Different forms of one are accepted, e.g. multiplying and dividing by the same number.	$6+2$ and $1 \cdot 6+2$
Factorization	A number is presented as a multiplication, i.e. the number has been factorized.	$8+4$ and $2 \cdot 4+4$
Using a quotient	A number is presented as a division.	$8+4$ and $16:2+4$

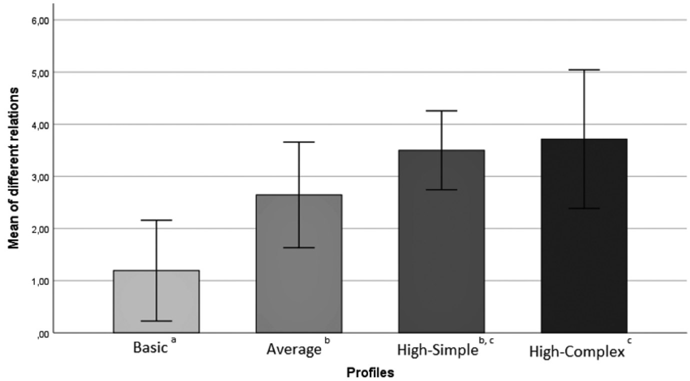

Figure 3. Mean values and standard deviations on the number of relations found in the students' answers for each profile

Note. Error bars = ± 1 SD. Statistically significant differences among the profiles are presented via indexes.

($SD = 1.31$), respectively. One-way ANOVAs showed that these differences were significant ($F(3, 200) = 61.36$, $p < .001$, $\eta^2 = .48$). Apart from these results, figure 3 presents the results from Tuckey's post-hoc test via indexes, which indicate that the students in the Basic profile had significantly fewer related solutions than those in the other profiles. There were also significant differences between the students in the Average and High-Complex profiles.

To determine whether the students changed their strategies to modify their solutions in different items or whether they used the same strategy throughout the test, we counted the number of items in which each relation appeared. Therefore, the value for each relation could range from 0 to 4. For instance, if a student used decomposition in every item, the value of the decomposition would be 4. Table 7 presents the mean values and standard deviations of every relation in each profile. Furthermore, the ANOVA test values are included in table 7, and the Tuckey's post-hoc test results are presented via indexes. As presented, for the students in the Basic profile, decomposition, using a difference, and repeated addition appeared significantly less frequently in the answers than in the other students' answers. Conversely, the students in the High-Simple profile had significantly higher mean values for decomposition and commutative law, compared to other students. For the High-Complex profile, there were significant differences with the Average profile in repeated addition and with the Basic profile in adding a zero. In addition, the students in the High-Complex profile had significantly higher mean values in factorization and using a quotient. However, there were no significant differences among the profiles in multiplying by one.

Table 7. *Means and standard deviations of the number of items in which each relation appeared, statistical values from one-way ANOVAs, and results from post-hoc tests for the profiles*

Relation	Basic	Average	High-Simple	High-Complex			
	M (SD)	M (SD)	M (SD)	M (SD)	F(3, 200)	*p*	η^2
Decomposition	$.51^{a}$ (.74)	1.66^{b} (.99)	2.75^{c} (.89)	1.15^{b} (1.06)	33.34	< .001	.33
Using a difference	$.15^{a}$ (.40)	$.47^{b}$ (.68)	$.75^{b}$ (.71)	$.67^{b}$ (.83)	9.09	< .001	.12
Commutative law	$.62^{a}$ (1.07)	1.00^{a} (1.35)	2.38^{b} (1.51)	$.78^{a}$ (1.19)	6.11	.001	.08
Repeated addition	$.24^{a}$ (.43)	$.73^{b}$ (.81)	$1.00^{b,c}$ (.93)	1.22^{c} (1.05)	19.67	< .001	.23
Adding a zero	$.04^{a}$ (.19)	$.17^{a,b}$ (.46)	$.50^{a,b}$ (1.41)	$.44^{b}$ (.84)	6.22	< .001	.09
Multiplying by one	$.04^{a}$ (.19)	$.03^{a}$ (.18)	$.00^{a}$ (.00)	$.15^{a}$ (.36)	2.14	.10	.03
Factorization	$.03^{a}$ (.16)	$.19^{a}$ (.47)	$.25^{a}$ (.46)	$.96^{b}$ (.81)	34.31	< .001	.34
Using a quotient	$.03^{a}$ (.16)	$.07^{a}$ (.25)	$.00^{a}$ (.00)	$.59^{b}$ (.69)	23.60	< .001	.26

Note. There are statistical differences only between the profiles that do not have the same letter in the index

Discussion

In this study, we found that students' adaptive number knowledge improved with grade level. This result supports the findings of previous studies (McMullen et al., 2017). Significant differences were found both in the simple and complex solutions; however, in the simple solutions in the sparse items, the differences were only between the second and sixth graders. In the complex solutions, there were significant differences among all grades in both the dense and sparse items. This validates the differences between dense and sparse items since there were not many easy simple solutions in the sparse items. Therefore, in order to have more solutions in the sparse items, the students had to produce complex solutions.

The grade-level differences arguably reveal the development of adaptive number knowledge throughout primary school. According to Blöte et al. (2000), the more a student gains experience and expertise in dealing with numbers, the more flexible their mental calculation becomes, and the easier it is for them to recognize relations between the numbers and apply strategies. Due to the lack of experience among second graders, the use of multiplication and division might be difficult for them because of the novelty of these operations to them (Opetushallitus, 2016). This might explain the large difference between the grades when analyzing the complex solutions in both item types. In addition, as students get older, they tend to prefer multiplication to addition (Van Dooren et al., 2010), so older students have a better chance of coming up with complex solutions. Another challenge for younger students is the time limit of the test; older students are able to perform addition faster and more accurately (Siegler & Shipley, 1995).

We found four latent profiles describing the students' adaptive number knowledge: Basic, Average, High-Simple, and High-Complex. In comparison to previous studies, there were some similarities and some differences between the profiles. McMullen et al. (2017) found five profiles for fourth to sixth graders: Basic, Simple, Complex, Strategic, and High. The Basic profiles in both studies are quite similar. In addition, the Simple profile in the study of McMullen et al. (2017) is similar to our High-Simple profile. The High-Complex profile in this study is a combination of the Complex profile, wherein students had many complex solutions and few simple solutions, especially in the sparse items, and the High profile, wherein all variables had high values. However, none of our profiles matched the Strategic profile of McMullen et al. (2017), wherein there were more simple solutions in the dense items and more complex solutions in the sparse items. Instead, we found the Average profile, wherein the students' solutions were rather close to the average (simple solutions above and complex solutions below). This may have resulted from the test modification. In the s-APT, the sparse items contained small numbers, which might have affected the students' solutions. Another reason for the differences between this and previous studies could be the larger age range and younger participants in the present study.

The eight relations emerging from the students' answers helped describe the differences among the profiles. An analysis of the number of relations and the number of items in which the relations appeared revealed significant differences among the profiles. Therefore, the related solutions elaborated the characteristics of the profiles and showed the various strategies used by the students to modify their solutions. According to previous studies (Heinze et al., 2009; Siegler & Lemaire, 1997; Verschaffel et al., 2009), students with flexible mathematical skills will choose their strategies in a versatile manner based on the problem. Therefore, as adaptive number knowledge is related to other mathematical skills such as arithmetic fluency (McMullen et al., 2017), we could presume that students with higher adaptive number knowledge will have various related solutions and might use different strategies in different tasks. The most common relation for the students in the Basic profile was commutative law, whereas for the Average and High-Simple profiles, the most common relation was decomposition. Furthermore, the students in the High-Complex profile favored repeated addition. It can be deduced from the high number of simple solutions, that the students frequently focused on the same number characteristics when producing their solutions (McMullen et al., 2017). The analysis of the related solutions supports this conclusion as the students in the High-Simple profile had substantially high mean values in decomposition and commutative law but rather low values in other relations, which means that they relied only on these two strategies and made no changes to their thinking in the various tasks.

Since the High-Complex profile had the highest number of complex solutions, it was not surprising that it contained significantly more solutions related by factorization and using a quotient, which require the use of multiplicative operations. Furthermore, the students in the High-Complex profile recorded the highest number of relations compared to all other profiles. In addition, their mean values in some relations were lower than in High-Simple profile, but overall, there was minimal variety in the mean values, which meant a diverse use of strategies. In other words, combining the results from figure 3 and table 7, we see that the students from the High-Complex profile used a range of strategies in a more versatile manner than other students. This supports the conclusion of previous studies (McMullen et al., 2017) that students in the High profile alternate their strategies in different items. Students' competence in choosing a suitable strategy helps them solve mathematical problems rapidly and accurately, but the converse does not apply (Hästö & Palkki, 2019). Therefore, students who used the most strategies might produce more solutions in a test, but a high number of correct solutions does not necessarily mean that the students have proficiently mastered different strategies. Nevertheless, not all pairs of solutions, even complex solutions, can be explained by any of the relations. Thus, some highly adaptive solutions might not be recognized through the relations.

The idea of the test is to use only a few numbers and basic arithmetic operations to reach the same target number. When performing calculations, it is also important to understand the characteristics of the operations used. According to Hatano (2003), adaptive expertise is the ability to understand the contents of the procedures and use their relations. This can be seen in students' answers, for example, if they produce solutions related by repeated addition, as $2 \cdot 3 = 6$ and $3 + 3 = 6$. Although repeated addition is considered a simple and primitive strategy (Van Dooren et al., 2010), answering these expressions can show that the student not only knows how to use the strategy, but also understands the relation between addition and multiplication. However, since the numbers that can be used are limited, the possible solutions are somewhat restricted. Therefore, as the number of correct solutions increases, the possibility of related solutions also increases even though the student may not consciously try to use them.

The item-by-item inspection of the answers showed that although we modified the sparse items to make them easier, the students still produced fewer correct solutions in the sparse than in the dense items. In addition, in the dense items, the students mainly produced more simple solutions and fewer complex solutions than in the sparse items, which is compatible with the results in previous studies (McMullen et al., 2017, 2019). This indicates that the modifications were reasonable and that the sparse items were not overly easy. However, if the students' solutions in the dense and sparse items were compared to those in previous studies, there would be a difference in the simple solutions in the sparse items. In McMullen et al. (2017), students from grades four to six produced substantially fewer simple solutions in the sparse items than those in this study, but the same phenomenon was not noted with complex solutions in the sparse items, neither with simple and complex solutions in the dense items. This explicit difference may be attributed to the modification made in the sparse items for the second graders. Overall, we found the same characteristics between the dense and sparse items as in previous studies (McMullen et al., 2017, 2019). Furthermore, similar characteristics between the profiles were found in our study as those from previous studies (McMullen et al., 2017, 2019). While a majority of the second and fourth graders were in the Basic profile, almost one-third of the sixth graders were also included. This and the significant differences among the grade levels suggest that adaptive number knowledge develops with age but that there is more variation among older students.

As the s-APT can identify individual differences by dividing the students into profiles, it can be used as a pre-test to distinguish the mathematical skills of primary school students of various ages. In the future, research on the associations between adaptive number knowledge and other mathematical or non-mathematical skills could be conducted with even younger students. In addition, the analysis of related solutions presents a new perspective to study how students recognize relations between numbers and solutions. The conscious use of these relations and their associations with other flexible mathematical skills could provide more insights into students' mathematical thinking.

Conclusion

We found significant differences between grade levels in both the quantity and quality of the students' answers. In addition, the students could be divided into four profiles based on their adaptive number knowledge. A new analysis of related solutions where the students' solutions were studied more thoroughly gave a better idea about the differences between the profiles. Related solutions provide indications about the strategies that students use to produce their solutions. This kind of analysis could be further developed and used to study other characteristics in students' solutions. For instance, including incorrect solutions in the analysis might also provide information about students' misconceptions in arithmetic. The results support previous studies, in that, there were both age-dependent and individual differences in the students' adaptive number knowledge. In addition, the findings suggest that the Arithmetic Production Task could be modified for use with younger students and that the modified version can identify differences in the students' adaptive number knowledge.

Acknowledgements

This work was funded by the Academy of Finland (project number 318010). We are grateful to Erno Lehtinen, Minna Hannula-Sormunen and the Mathematical learning research group at the University of Turku for providing materials for implementing the *Arithmetic production task*.

References

Berlin, K. S., Williams, N. A. & Parra, G. R. (2014). An introduction to latent variable mixture modeling (part 1): overview and cross-sectional latent class and latent profile analyses. *Journal of Pediatric Psychology*, 39 (2), 174–187. doi: 10.1093/jpepsy/jst084

Blöte, A. W., Klein, A. S. & Beishuizen, M. (2000). Mental computation and conceptual understanding. *Learning and Instruction*, 10 (3), 221–247. doi: 10.1016/S0959-4752(99)00028-6

Brezovszky, B., McMullen, J., Veermans, K., Hannula-Sormunen, M. M., Rodríguez-Aflecht, G. et al. (2019). Effects of a mathematics game-based learning environment on primary school students' adaptive number knowledge. *Computers & Education*, 128, 63–74. doi: 10.1016/j.compedu.2018.09.011

Brezovszky, B., Rodrígues-Aflecht, G., McMullen, J., Veermans, K., Pongsakdi, N. et al. (2015). Developing adaptive number knowledge with the number navigation game-based learning environment. In J. Torbeyns, E. Lehtinen & J. Elen (Eds.), *Describing and studying domain-specific serious games* (pp. 155–170). Springer. doi: 10.1007/978-3-319-20276-1_10

Gravemeijer, K., Bruin-Muurling, G., Kraemer, J.-M. & Stiphout, I. van (2016). Shortcomings of mathematics education reform in the netherlands: A paradigm case? *Mathematical Thinking and Learning*, 18 (1), 25–44. doi: 10.1080/10986065.2016.1107821

Hästö, P. & Palkki, R. (2019). Finnish students' flexibility and its relation to speed and accuracy in equation solving. *Nordic Studies in Mathematics Education*, 24 (3-4), 43–58.

Hatano, G. (2003). Foreword. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: constructing adaptive expertise* (pp. xi–xiii). Lawrence Erlbaum.

Heinze, A., Star, J. R. & Verschaffel, L. (2009). Flexible and adaptive use of strategies and representations in mathematics education. *ZDM*, 41 (5), 535–540. doi: 10.1007/s11858-009-0214-4

Krippendorff, K. (2004). Reliability in content analysis some common misconceptions and recommendations. *Human Communication Research*, 30 (3), 411–433. doi: 10.1111/j.1468-2958.2004.tb00738.x

Lemaire, P. & Siegler, R. S. (1995). Four aspects of strategic change: contributions to children's learning of multiplication. *Journal of Experimental Psychology: General*, 124 (1), 83–97. doi: 10.1037/0096-3445.124.1.83

McMullen, J., Brezovszky, B., Hannula-Sormunen, M. M., Veermans, K., Rodríguez-Aflecht, G. et al. (2017). Adaptive number knowledge and its relation to arithmetic and pre-algebra knowledge. *Learning and Instruction*, 49, 178–187. doi: 10.1016/j.learninstruc.2017.02.001

McMullen, J., Brezovszky, B., Rodríguez-Aflecht, G., Pongsakdi, N., Hannula-Sormunen, M. M. & Lehtinen, E. (2016). Adaptive number knowledge: exploring the foundations of adaptivity with whole-number arithmetic. *Learning and Individual Differences*, 47, 172–181. doi: 10.1016/j.lindif.2016.02.007

McMullen, J., Kanerva, K., Lehtinen, E., Hannula-Sormunen, M. M. & Kiuru, N. (2019). Adaptive number knowledge in secondary school students: profiles and antecedents. *Journal of Numerical Cognition*, 5 (3), 283–300. doi: 10.5964/jnc.v5i3.201

Muthén, L. K. & Muthén, B. O. (2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.

Nylund, K. L., Asparouhov, T. & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: a monte carlo simulation study. *Structural Equation Modeling: a Multidisciplinary Journal*, 14 (4), 535–569. doi: 10.1080/10705510701575396

Opetushallitus (2016). *National core curriculum for basic education 2014*. Finnish National Board of Education. https://www.oph.fi/en/statistics-and-publications/publications/new-national-core-curriculum-basic-education-focus-school

Selter, C. (2009). Creativity, flexibility, adaptivity, and strategy use in mathematics. *ZDM*, 41 (5), 619–625. doi: 10.1007/s11858-009-0203-7

Siegler, R. S. & Lemaire, P. (1997). Older and younger adults' strategy choices in multiplication: testing predictions of ASCM using the choice/no-choice method. *Journal of Experimental Psychology: General*, 126(1), 71–92. doi: 10.1037/0096-3445.126.1.71

Siegler, R. S. & Shipley, C. (1995). Variation, selection, and cognitive change. In T. J. Simon & G. S. Halford (Eds.), *Developing cognitive competence: new approaches to process modeling* (pp. 31–76). Psychology Press.
Star, J. R. & Newton, K. J. (2009). The nature and development of experts' strategy flexibility for solving equations. *ZDM*, 41(5), 557–567. doi: 10.1007/s11858-009-0185-5
Tein, J.-Y., Coxe, S. & Cham, H. (2013). Statistical power to detect the correct number of classes in latent profile analysis. *Structural Equation Modeling: a Multidisciplinary Journal*, 20 (4), 640–657. doi: 10.1080/10705511.2013.824781
Threlfall, J. (2002). Flexible mental calculation. *Educational Studies in Mathematics*, 50 (1), 29–47. doi; 10.1023/A:1020572803437
Van Dooren, W., De Bock, D. & Verschaffel, L. (2010). From addition to multiplication ... and back: the development of students' additive and multiplicative reasoning skills. *Cognition and Instruction*, 28 (3), 360–381. doi: 10.1080/07370008.2010.488306
Verschaffel, L., Luwel, K., Torbeyns, J. & Van Dooren, W. (2009). Conceptualizing, investigating, and enhancing adaptive expertise in elementary mathematics education. *European Journal of Psychology of Education*, 24 (3), 335–359. doi: 10.1007/BF03174765

Salla Pehkonen

Salla Pehkonen (MSc) is a PhD student at the University of Jyväskylä. Her research focuses on students' adaptive mathematical skills and their role in implicitly mathematical situations.

salla.m.pehkonen@jyu.fi

Antti Lehtinen

Antti Lehtinen (Ph.D.) is a senior lecturer at the University of Jyväskylä His research encompasses technology-enhanced inquiry-based teaching and learning, the development of university-level physics instructional labs, differentiation in science teaching, and classroom interaction.

antti.t.lehtinen@jyu.fi

Pasi Nieminen

Pasi Nieminen (Ph.D.) is a senior lecturer at the University of Jyväskylä. His research areas include inquiry-based science and mathematics learning from different perspectives such as formative assessment, argumentation, teacher guidance, and differentiation.

pasi.k.nieminen@jyu.fi

Markus Hähkiöniemi

Markus Hähkiöniemi (Ph.D., docent) is a senior lecturer at the University of Jyväskylä. His research interests include technology-enhanced inquiry-based mathematics teaching, argumentation, and classroom interaction.

markus.hahkioniemi@jyu.fi