Finnish students' flexibility and its relation to speed and accuracy in equation solving

Peter Hästö and Riikka Palkki

PH: University of Turku and University of Oulu, Finland; peter.hasto@oulu.fi;

RP: University of Oulu, Finland; riikka.palkki@oulu.fi

A total of 266 Finnish students participated in a flexible equation solving test. By flexibility we understand the knowledge of multiple strategies and ability to choose the most mathematically appropriate strategy for a given task. Here we focus on the first aspect, namely knowledge of appropriate alternative, so-called innovative strategies. The test measured students' capacity and inclination for producing innovative strategies. We consider the relationship between these measures and students' speed and accuracy in solving equations. We find that students with high capacity for innovation have high speed and accuracy. On the other hand, some low capacity students had high speed or accuracy whereas others had low. Inclination for innovation is not related to speed or accuracy.

Keywords: Equation solving, flexibility, adaptive expertise, secondary-school, algebra, innovative strategies.

Introduction

Mathematical tasks can often be solved in multiple ways. Skilled problem solvers know multiple strategies and can use them effectively in various situations. It has been shown that comparing multiple strategies is a powerful tool to improve students' learning of mathematics (see Star et al., 2015, for a review). However, many students believe that mathematical tasks can be solved in only one correct way which has been shown in advance by the teacher (Schoenfeld, 1992).

Linear equation solving is an area where every task can be solved by a single procedure, known as the standard strategy (cf. Table 2 on page 4). Nevertheless, equations are a challenge for students and errors persist when solving equations (Booth, Barbieri, Eyer, & Paré-Blagoev, 2014). In the Finnish education system, algebra is seen as an area where the focus should be on learning rules for manipulating algebraic expressions and equations, i.e. procedural knowledge (Attorps, 2006; Hihnala, 2005; see also Andrews, 2013, on more recent conceptual emphasis).

Star (2005) criticized the view that procedural and conceptual knowledge correspond to superficial and deep understanding, respectively; he posited that flexibility is a form of deep procedural knowledge. In the field of learning arithmetic, the term "adaptive expertise" has been used to describe the same phenomenon (e.g. McMullen et al., 2017; Torbeyns, Verschaffel & Ghesquière, 2006; Verschaffel, Greer & De Corte, 2007). Flexibility is a step from mechanical use of fixed rules towards more reflective modes of working and requires both conceptual and procedural knowledge (Schneider, Rittle-Johnson, & Star, 2011).

While flexibility is explicitly emphasized in many curriculum documents (e.g. NCTM, 2000), the Finnish national standards refer only to related concepts such as creativity and allowing each student to develop their own strategies (Opetushallitus, 2014). In our work with dozens of Finnish teachers, we have found that there are mixed feelings about the importance of flexibility in school mathematics. Similarly, Buchbinder, Chazan and Fleming (2015) found that US teachers' emphasis on the standard strategy in equation solving effectively hinders them from using multiple solutions to develop flexibility. On the other hand, Finnish mathematics teachers saw many positive aspects in using comparison of multiple solutions in learning mathematics (Palkki, 2018).

Obtaining a correct answer quickly is often seen by students as the paramount indicator of mathematical proficiency; mathematics education researchers, on the other hand, emphasize conceptual understanding, problem-solving skills and adaptive expertise (Hatano, 2003; Schoenfeld, 1992; Törner, 1998). Furthermore, many students think committing errors is catastrophic whereas educators see great value also in mistakes (Streuer & Dresel, 2015). Thus, students may see mathematics as a "time contest". In elementary grades, Lemaire and Siegler (1995) found in that improved adaptivity was one source of increased speed and accuracy in multiplication tasks. On the other hand, Verschaffel, Greer and De Corte (2007) reviewed multiple studies to conclude that most children lack the disposition to use adaptive mental arithmetic when a well-learned standard strategy worked, even on tasks like $4002-3998$ where they would provide great benefit.

It seems that less attention has been paid to the relationship between flexibility and speed and accuracy in middle- and high-school mathematics. In this article we compare such students' performance on measures of speed and accuracy (highly valued by students) with their flexibility (highly valued by mathematics educators). We next describe these terms more precisely and state our research hypotheses.

The *strategy* of an equation-solving task is the chain of deductions (intermediate steps) required to reach the answer. The strategy will be deduced from the solution given by the student to the task as described in the next section.[1] We call the ability to correctly carry out a strategy's individual steps *accuracy* whereas *speed* refers to the number of tasks solved in a given timespan. Flexibility (more precisely strategic flexibility) is related to choosing the strategy which determines what individual steps are needed. A flexible problem solver knows different approaches and is able to choose the most mathematically appropriate one (Rittle-Johnson & Star, 2007; Star & Rittle-Johnson, 2008). In this article, we focus on the first aspect of flexibility, namely knowledge of appropriate alternative. In the test used in this study, from Xu, Liu, Star, Wang, Liu and Zhen (2017), the efficient approach was always a so-called innovative strategy (cf. Table 2), hence we use the term *innovative* in what follows. In our test, participants could use an innovative strategy either spontaneously or as a result of being prompted for multiple solutions. We took an innovative strategy produced in either case as evidence of *capacity* (for innovation). By *inclination* (towards innovation), on the other hand, we mean a predilection for spontaneous innovative strategy. For the operationalization of these concepts, see the section "Variables", below.

Many students may neither have been provided with the opportunity to learn flexibility in equation solving nor see the value in them. On the other hand, students who are quick and accurate may create such opportunities for themselves by solving given tasks in non-standard or multiple ways. Of course, quick and accurate students may also not do so. Hence, we expect both capacity and inclination to be related to speed and accuracy. To be more specific, our hypotheses are that

1. higher capacity for innovation is related to higher speed and higher accuracy, and
2. higher inclination toward innovation is related to higher speed and higher accuracy.

[1] Note that by "solution" we always refer to the writing produced by the student on the test-sheet. In the context of equations, "solution" may refer to the values of the variable which makes the equation true. In this article, this sense of the word "solution" is never used.

There is also a connection in the opposite direction: flexibility enables solving tasks quickly and correctly (Heinze, Star, & Verschaffel, 2009; Lemaire & Siegler, 1995). We will consider this in the discussion.

Method

Context

Finnish children start school at age 7 and attend compulsory school until grade 9 with a single compulsory mathematics curriculum for everyone. Linear algebraic equations appear in 7th or 8th grade in full generality. High school (grades 10–12) is not compulsory and about half of the age group attends. Most of the remaining students attend vocational school. In high school students choose between two tracks in mathematics: advanced and basic.

In TIMSS, algebra was a weakness for Finnish students (Kupari, Vettenranta & Nissinen, 2012). In contrast, they have fared well in the PISA test, which involves, among others, the ability to apply simple equations in context (Kupari et al., 2013).

Participants

We collected data on the ability to produce standard and innovative strategies in linear equation solving with the Tri-Phase Flexibility Assessment test (cf. “Procedure”, below) and a convenience sample. The participating schools were diverse in size and geographic location (both rural and urban).

A total of 266 students participated in the study. The data consist of 93 tests from 8th graders and 164 tests from 11th graders of which 103 studied advanced mathematics and 61 studied basic mathematics. Also included were 9 students in technical vocational school.

Of the 8th grade students, 75 had studied to some extent with material geared toward conceptual understanding of equation solving and flexibility produced by our research group for another project.

Procedure

Students completed a 12-equation test, the Tri-Phase Flexibility Assessment, during a regular 45-minute class. The equations used in our study are shown in table 1. In the first phase of the test (15 min), students were asked to produce one solution to each of the 12 equations. In the second phase (20 min), students were asked to produce as many solutions as possible for the same 12 equations. In the third phase (5 min), students were asked to choose “the best” solution for each of the tasks. The test was devised by Xu et al. (2017), who established its psychometric reliability and validity. It has also been used by Liu et al. (2018) to study mediating effects of beliefs on the relation between potential and practical flexibility; by Joglar Prieto, Abánades Astudillo and Star (2018) to study flexibility in the Spanish school system; and by Hästö, Palkki, Tuomela and Star (2019) to examine the relationship between flexibility and matriculation examination results.

Table 1. The equations in the test

1) $4(x-2)=24$	2) $3(x+0{,}69)=15$	3) $4(x+\frac{3}{5})=12$
4) $4(x+6)+3(x+6)=21$	5) $5(x+\frac{3}{7})+3(x+\frac{3}{7})=16$	6) $2(x-0{,}31)+3(x-0{,}31)=15$
7) $8(x-5)=3(x-5)+20$	8) $8\left(x-\frac{2}{5}\right)-11=6(x-\frac{2}{5})$	9) $5(x+0{,}6)+3x=5(x+0{,}6)+7$
10) $\frac{2x-6}{2}+\frac{6x-18}{3}=5$	11) $\frac{x+3}{3}+\frac{3x-9}{9}=1$	12) $\frac{5x+5}{5}+\frac{6x+6}{6}=6$

Every linear equation in one variable can be solved with the standard strategy. It consists of the steps *distribute-parentheses; combine-like-terms; move x-terms to the left and non-x-terms to the right; divide-by-coefficient*. It is commonly taught in Finnish schools. The equations in the test can be solved by the standard strategy, but each task also allowed for an "innovative" strategy. As an example, the standard and innovative strategies of the fourth equation are shown in Table 2. The innovative strategy for each of the 12 equations is given by Xu et al. (2017) and is part of the Tri-Phase Flexibility Assessment's scoring manual.

Table 2. Standard and innovative strategies

Standard strategy	Innovative strategy
$4(x+6)+3(x+6)=21$	$4(x+6)+3(x+6)=21$
$4x+24+3x+18=21$	$7(x+6)=21$
$7x+42=21$	$x+6=3$
$7x=-21$	$x=-3$
$x=-3$	

Coding

The solutions were coded following the scheme of Xu et al. (2017). Each solution was coded for correctness and strategy. A solution was marked as correct if all steps and the answer were correct. Three types of strategies were distinguished in the coding: standard, innovative, and other. Strategy was assessed independent of correctness. Standard strategy means following the general procedure detailed above, whereas innovative strategies were specified in advance for each task (see Xu et al., 2017). Only the first step (i.e. $7(x+6)=21$ in the example above) was required for a solution to be coded as innovative. All other strategies were in the category "other".

Writing from which the intended strategy could not be discerned (e.g. only an answer is given, or some gibberish is written) were not registered, i.e. they do not count as a solution or an attempt.

Coding was done by a research assistant. Twelve students' tasks were independently recoded by two other members of our research group. The inter-rater reliability was 98%, which is considered very high.

Variables

As a measure of student *speed*, we used the number of tasks with a solution in the first phase of the test, regardless of correctness.

It was challenging to disentangle accuracy from speed, since the former must somehow also rely on tasks solved correctly, and you cannot solve a task correctly if you have not solved it at all. For this reason, we used only equations 1–9 to determine accuracy, since equations 10–12 were solved only by few students and using them would lead to a strong connection with speed.

Accuracy means the number of tasks solved correctly divided by sum of the number of tasks solved correctly and the number of tasks solved incorrectly. Note that a task may contain both a correct and an incorrect solution.

A student's *capacity* was defined as the number of tasks to which the student produced an innovative strategy in either the first or the second phase. A student's *inclination* was the number of tasks to which the student produced an innovative strategy in the first phase (i.e. without extra prompting) less the number of tasks to which the student produced in innovative strategy in the second phase. Thus, positive inclination means a tendency to offer the innovative strategy unprompted, negative inclination a tendency to use other strategies (usually the standard one) in the first phase while zero inclination means a no preference between these two.

The variables used in this study are summarized in Table 3.

Table 3. Variables

Variable (abbreviation)	**Definition**
capacity (cap)	The number of tasks with an innovative strategy
inclination (incl)	The number of tasks in first phase with an innovative strategy less the number of tasks in later phase with an innovative strategy
speed (speed)	The number of tasks attempted in the first phase of the test, regardless of correctness
accuracy (acc)	The number of tasks solved correctly divided by sum of the numbers of tasks solved correctly and tasks solved incorrectly

Let us digress to discuss the measures of flexibility used in this study and earlier studies based on the Tri-Phase Flexibility Assessment. Xu et al. (2017) as well as Liu et al. (2018) coded solutions for correctness and strategy as described above. However, instead of capacity and inclination, they used the variables practical flexibility and potential flexibility. A task is scored for practical flexibility if the strategy in the first phase is innovative. A task is scored for potential flexibility if it has a standard strategy and an innovative strategy and the innovative strategy is marked as best. The Tri-Phase Flexibility Assessment has also been used with different flexibility variables by Joglar Prieto, Abánades Astudillo and Star (2018) and Hästö, Palkki, Tuomela and Star (2019).

Note that the practical flexibility and potential flexibility variables are partially based on the same criterion, namely the presence of an innovative strategy. In the Chinese sample of Liu et al. (2018) the correlation between the variables was 0.30. However, a preliminary investigation of the Finnish data revealed a higher correlation; this is because almost all Finnish students chose the innovative strategy as best, so that the relation "practical flexibility ≤ potential flexibility" holds in the Finnish data, which forces a strong correlation. Note that capacity measures the main criterion of potential flexibility (presence of an innovative strategies) whereas (capacity+inclination)/2 gives the number of innovative strategies in the first phase of the test, i.e. it equals the practical flexibility variable.

Analysis and results

For a general picture of the test results, we start with some descriptive statistics. First, we summarize the number of students and groups in this study.

Table 4. Number of students and groups

Group	Diagram color	# of classes	# of students
8th grade	blue	6	93
basic math, 11th grade	red	4	61
advanced math, 11th grade	green	5	103

The number of tasks with an innovative strategy ranged from 0 to 12, i.e. the whole range of possible values. Students produced at least one solution to a total of 2427 tasks, an average of 9.12 solved tasks per student (out of 12). The total number of tasks with at least one innovative strategy was 853, an average of 3.2 tasks per student (out of 12). Of these innovative strategies, 734 were marked as the best strategy in the third phase, whereas in the remaining 119 tasks some other strategy (or none) was marked as best. Table 5 shows the distribution of innovative strategies based on the phase of the rest in which they were produced and whether they were marked as best or not.

Table 5. The number of tasks with innovative strategies

Innovative strategies in the…	Marked as best	Not marked as best	Total
… first phase	339	21	360
… later phases	395	98	493
Total	734	119	853

Figure 1 shows the number of correct (bright color) and incorrect (dark color) tasks. Tasks are divided into four groups of three tasks based on their structure (cf. Table 1) and the colors correspond to the three groups of students (Table 4). For instance, from the first green bar we see that students in advanced mathematics provided an average of about 1.75 correct solutions per task for each of tasks 1–3 and an additional 0.25 incorrect solutions per task for the same tasks. When red (basic mathematics) and green (advanced mathematics) charts are compared, it is seen that students in advanced mathematics have many more correct solutions than students in basic mathematics. Furthermore, students in general did fewer tasks toward the end of the test.

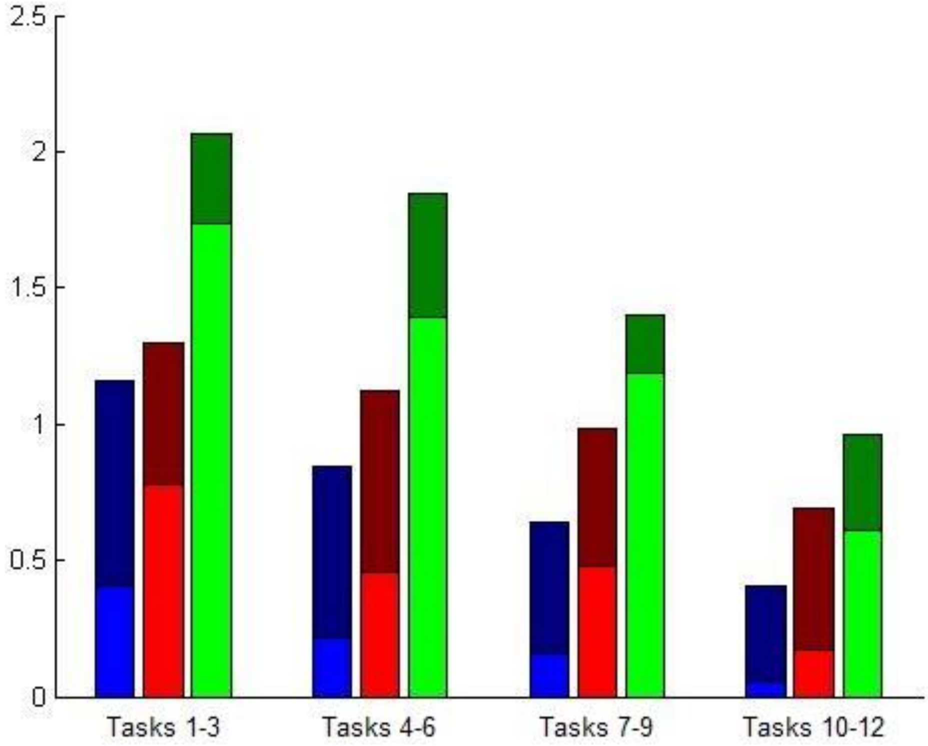

Figure 1. Students' average number of correct (bright) and incorrect (dark) solutions per task in different task groups

The task groups and colors in Figure 2 are the same as in Figure 1, but this time we compare the average amount of innovative and standard strategies per task. We see that students in basic mathematics had very few innovative strategies, even fewer than students in 8th grade even though they had many more solutions in total.

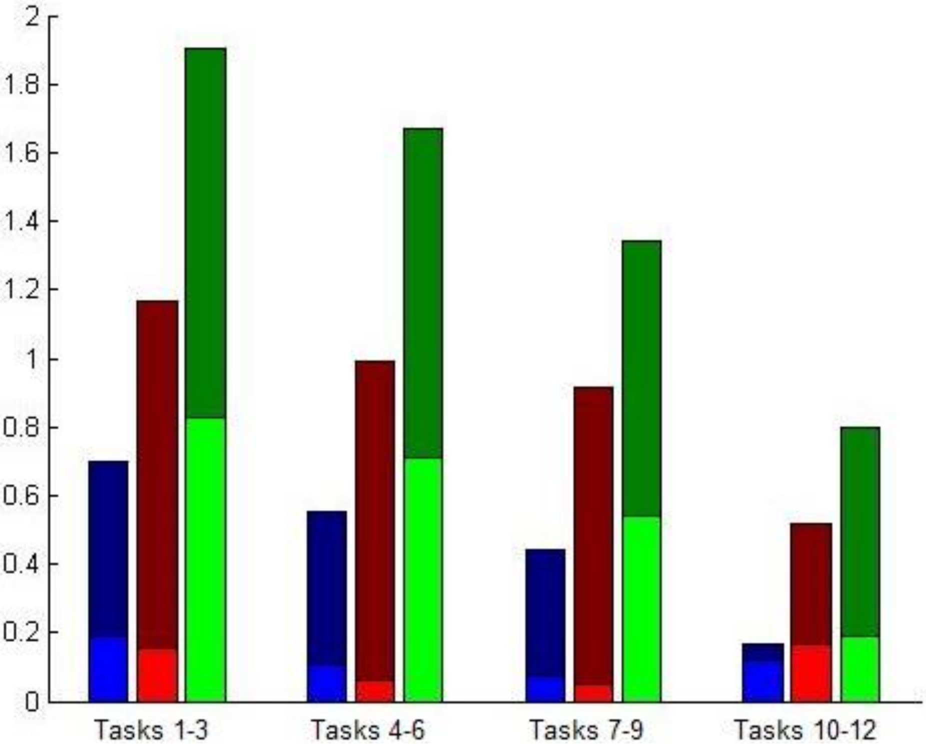

Figure 2. Students' average number of innovative (bright) and standard (dark) strategies per task in different task groups

Let us move on to the statistical tests. We analyzed the bivariate Pearson correlations between the speed and accuracy variables to determine to what extent they are independent and tested whether their correlations differ statistically significantly from zero. The correlation between speed and accuracy was 0.48 ($p < 0.001$). Accuracy was quite highly correlated with speed, and therefore not optimally suited to measure skill independent of speed.

The same correlation analysis was performed also for capacity and inclination. The correlation between them was -0.06 ($p > 0.1$). Capacity and inclination were essentially uncorrelated, which indicates that they may capture different aspects of students' mathematical behavior, and that these measures form a reasonable basis for regression analysis.

Next, we used linear regression analyses (Freund, Wilson, & Sa, 2006) with capacity and inclination as independent variables and speed and accuracy as dependent variables. This choice of which variable set to use as independent variables is based on the fact that speed and accuracy were linearly dependent and would thus not be reliable independent variables in a regression analysis due to multicollinearity. The analyses were carried out with SPSS version 22, correlation was measured as bivariate and linear regression was performed with the "enter" method and both independent variables in the same block.

The results of the linear regressions are shown in Table 6. Capacity is a statistically highly significant predictor for each dependent variable with $p < 0.0005$. As can be seen from the high p-values, inclination was not a statistically significant predictor of speed or accuracy.

Table 6. The results of the linear regression analyses ($n = 266$)

Dependent variable	Independent variable	Standardized β	t	p
Speed	capacity	$.535$	10.280	$.000$
$R^2 = .29, F = 53.4, p < 0.001$	inclination	$-.024$	$-.454$	$.650$
Accuracy	capacity	$.671$	14.650	$.000$
$R^2 = .45, F = 107.7, p < 0.001$	inclination	$-.003$	-0.073	$.942$

Recall that the square of the standardized beta coefficient indicates what extent of the variance of the dependent variable is explained by the corresponding independent variable in this data. Regression analyses showed that capacity has quite a strong relationship to speed and accuracy, explaining 29% and 45% of the variation, respectively. Whether a student is inclined to use innovative strategies or not did not have a statistically significant co-variation with speed or accuracy.

Table 7a. Cross-tabulation of speed and capacity

		Speed												
		1	2	3	4	5	6	7	8	9	10	11	12	Total
Capacity	0	4	6	6	9	13	11	11	17	5	11	7	4	104
	1			1	2		3	3	4	3	3	1	4	24
	2			1		3	4	1	3	3	3	1	1	20
	3			1	1	2	1	3	1	2	6	3	6	26
	4								1	1		1	4	7
	5						1	1	2	1	2	1	4	12
	6						1	1	2	1	1	1	4	11
	7								2	1	1	1	3	8
	8						1	2	4		3	2	5	17
	9								1	2	4	1	9	17
	10										2	2	7	11
	11								1				5	6
	12												3	3
Total		4	6	9	12	18	22	22	38	19	36	21	59	266

To obtain additional insight into the relationship between the variables, we performed cross-tabulations between capacity and speed and accuracy, i.e. the pairs which showed statistically significant relationships in the regression analyses. For speed, the complete cross-tabulation is shown in Table 7a. Recall that the table means, e.g., that there were 13 students with value 5 for speed and 0 for capacity, etc. Empty cells mean that there are no such students.

Table 7b. Summary of cross-tabulation of speed and capacity

	low speed (1-6)	high speed (7-12)	Total
low cap (0-6)	70	134	204
high cap (7-12)	1	61	62
Total	71	195	266

In Tables 7b and 8 we present summary tables of the cross-tabulations between capacity and speed and accuracy. In all summary tables the ranges have been split at the middle into just two categories, which are labeled "high" and "low".

Table 8. Summary of cross-tabulation of accuracy and capacity

	low acc (0-0.5)	high acc (0.5-1.0)	Total
low cap (0-6)	112	92	204
high cap (7-12)	1	61	62
Total	113	153	266

From Table 7a we see that the relationship "speed ≥ capacity" holds for 261 of the 266 students. In Tables 7a, 8 and 9 we note that there is in each case one cell with only a single student, who has high capacity and low speed or low accuracy.

Therefore, we see that in this sample, high capacity is related to high speed. Equivalently low speed is related to low capacity, but not the other way around, i.e. there are plenty of students with high speed and low capacity. Similarly, high capacity is related to high accuracy, but not the other way around (Table 8).

The research hypothesis 1 was thus borne out by the data, whereas hypothesis 2 was mostly not confirmed: inclination played almost no role in determining speed and accuracy.

Discussion

As expected, high capacity was related to high accuracy and high speed. On the other hand, students' inclination was not related to accuracy and speed, contrary to our expectation. This casts some doubts on the hypothesis that student predilection for innovative strategies is a vehicle for development of flexibility in a skills-oriented classroom.

However, it is worth noting that our measure of inclination measures what kind of strategies students present in a test-taking situation. It is conceivable that many students believe that the standard strategy is what is expected of them and thus offered it first (cf. the dispositional obstacles to flexibility discussed by Verschaffel, Greer and De Corte, 2007). Also, it is known that students do not necessary use more effective strategies even when they know them (e.g. Torbeyns et al., 2006). After it became clear in the second phase that the test was about more than the standard strategy, 86 % (Table 5) marked the innovative strategy is best, even when it was not the first strategy they presented. This supports the hypothesis that students might employ the innovative strategies more in a non-test-taking situation.

We also found that high speed and accuracy is not related to high levels of capacity. This is consistent with reform-oriented approaches which stress that learning basic competences does not lead to higher level proficiency (NCTM, 2000). It also fits with the finding by Torbeyns et al. (2006) that both below average and above average students displayed limited flexibility, whereas top-students are more (spontaneously) flexible: we expect both above average and top students to be able to complete these equations with accuracy and speed, at least in high school, but only the latter are flexible according.

Some caveats should be pointed out concerning the limitations of this study. First, students answered only one test, and we extracted all our measures from the data set. Therefore, the same piece of data may be counted in different variables, thus potentially introducing artificial correlation. As explained above, we tried to limit adverse effects by measuring accuracy only with solutions from tasks 1–9 rather than all tasks. In addition, as pointed out by Lemaire and Siegler (1995), both adaptive strategy choice and efficient strategy execution lead to increased accuracy and speed; in particular, as can be seen in Table 2, the innovative strategies in this study consist of fewer steps, each of which are technically less demanding. Therefore, our variables for accuracy and speed are influenced not only by efficient implementation of the strategy, but also by strategy choice. The effect of this is limited, but not eliminated, by the fact that most innovative strategies were produced in the second phase of the test. Furthermore, a relative measure of accuracy was used to limit the correlation with speed. However, these efforts were only partially successful. A more accurate picture could be obtained by evaluating student speed and accuracy in a separate test where clever strategic choices are not available.

Our study was cross-sectional, so it is not possible to determine whether flexibility allowed students to develop greater speed and accuracy, or the other way around, or whether both are caused by some additional factor not considered in this study. However, the 8th grade students showed a range of different speeds and accuracy but very few used innovative strategies, which suggests that flexibility comes after accuracy. Unfortunately, we are no closer to answer perhaps the most interesting question in this area: what causes some quick and accurate skilled students to become flexible equation solvers whereas others show no hint of flexibility; this remains an issue for future research.

To summarize, we found that high capacity for innovation was related to high accuracy and speed, whereas students with low capacity innovation could have either high and low accuracy and speed. Our most significant finding is that students' inclination had no relation with their accuracy and speed.

Acknowledgements

We would like to thank Jon R. Star for sharing the test employed in this study and Anna-Helena Isopahkala for coding the results. We would also like to thank the aforementioned as well as Miguel Abánades Astudillo, Nuria Joglar Prieto and Dimitri Tuomela for discussions about the test and measures used in this study. We are grateful to the anonymous referees for comments and suggestions. The second author was supported financially by the Jenny and Antti Wihuri Foundation.

References

Andrews, P. (2013). Finnish mathematics teaching from a reform perspective: A video-based case-study analysis. *Comparative Education Review*, 57(2), 189–211. DOI: 10.1086/669124

Attorps, I. (2006). *Mathematics teachers' conceptions about equations*. (Doctoral dissertation). Helsinki: University of Helsinki.

Booth, J. L., Barbieri, C., Eyer, F., & Paré-Blagoev, E. J. (2014). Persistent and pernicious errors in algebraic problem solving. *Journal of Problem Solving*, 7, 10–23. DOI: 10.7771/1932-6246.1161

Buchbinder, O., Chazan, D., & Fleming, E. (2015). Insights into the school mathematics tradition from solving linear equations. *For the Learning of Mathematics*, 35, 2–8.

Freund, R. J., Wilson, W. J., & Sa, P. (2006). *Regression analysis*. Burlington: Academic Press.

Hästö, P., Palkki, R., Tuomela, D., & Star, J. (2019). Relationship between mathematical flexibility and success in national examinations. *European Journal of Science and Mathematics Education*, 7, 1–13.

Hatano, G. (2003). Foreword. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills*, 11-14. Mahwah: Erlbaum.

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

Hihnala, K. (2005). *Laskutehtävien suorittamisesta käsitteiden ymmärtämiseen. Peruskoululaisen matemaattisen ajattelun kehittyminen aritmetiikasta algebraan siirryttäessä* [Transition from the performing of arithmetic tasks to the understanding of concepts. The development of pupils' mathematical thinking when shifting from arithmetic to algebra in comprehensive school]. (Doctoral dissertation). University of Jyväskylä.

Joglar Prieto, N., Abánades Astudillo, M. Á., & Star, J. (2018). Flexibilidad matemática y resolución de ecuaciones lineales [Mathematical flexibility and linear equation solving]. *Uno: Revista de Didáctica de las Matemáticas*, 080, 51–57.

Kupari, P., Vettenranta, J., & Nissinen, K. (2012). *Oppijalähtöistä pedagogiikkaa etsimään: kahdeksannen luokan oppilaiden matematiikan ja luonnontieteiden osaaminen: kansainvälinen TIMSS-tutkimus Suomessa* [Looking for student centered pedagogy: eight grade students’ mathematics and science proficiency: international TIMSS-study in Finland]. Jyväskylä: Koulutuksen tutkimuslaitos.

Kupari, P., Välijärvi, J., Andersson, L., Arffman, I., Nissinen, K., Puhakka, E., & Vettenranta, J. (2013). *PISA12 ensituloksia* [PISA12 first results]. Jyväskylä: Opetus- ja kulttuuriministeriö.

Lemaire, P., & Siegler, R. S. (1995). Four aspects of strategic change: Contributions to children’s learning of multiplication. *Journal for Experimental Psychology: General*, 124, 83–97.

Liu, R.D., Wang, J., Star, J. R., Zhen, R., Jiang, R. H., & Fu, X. C. (2018). Turning potential flexibility into flexible performance: Moderating effect of self-efficacy and use of flexible cognition. *Frontiers in psychology*, 9, 646. DOI: 10.3389/fpsyg.2018.00646

McMullen, J., Brezovszky, B., Hannula-Sormunen, M.M., Veermans, K., Rodríguez-Aflecht, G., Pongsakdi, N., & Lehtinen, E. (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

NCTM. (2000). Principles and standards for school mathematics. Reston: NCTM.

Opetushallitus. (2014). Perusopetuksen opetussuunnitelman perusteet 2014 [National core curriculum for basic education 2014]. *Opetushallitus, määräykset ja ohjeet 2014: 96*. Tampere: Juvenes Print – Suomen yliopistopaino Oy.

Palkki, R. (2018). Matematiikan opettajien ja opettajaopiskelijoiden käsityksiä vertailumenetelmästä. LUMAT: *International Journal on Math, Science and Technology Education*, 6(1), 105–128. DOI: 10.31129/LUMAT.6.1.327

Rittle-Johnson, B., & Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations. *Journal of Educational Psychology*, 99(3), 561–574. DOI: 10.1037/0022-0663.99.3.561

Schneider, M., Rittle-Johnson, B., & Star, J. R. (2011). Relations among conceptual knowledge, procedural knowledge, and procedural flexibility in two samples differing in prior knowledge. *Developmental Psychology,* 47(6), 1525–1538. DOI: 10.1037/a0024997

Schoenfeld, A. H. (1992). Learning to think mathematically: Sense-making in mathematics. In D. Grouws (Ed.). *Handbook for research on mathematics teaching and learning*, 334-370. New York: MacMillan.

Star, J. R. (2005). Reconceptualizing Procedural Knowledge. *Journal for Research in Mathematics Education*, 36, 404–411. DOI: 10.2307/30034943

Star, J. R., Pollack, C., Durkin, K., Rittle-Johnson, B., Lynch, K., Newton, K., & Gogolen, C. (2015). Learning from comparison in algebra. *Contemporary Educational Psychology,* 40, 41–54. DOI: 10.1016/j.cedpsych.2014.05.005

Star, J. R., & Rittle-Johnson, B. (2008). Flexibility in problem solving: The case of equation solving. *Learning and Instruction,* 18(6), 565–579. DOI: 10.1016/j.learninstruc.2007.09.018

Steuer, G., & Dresel, M. (2015). A constructive error climate as an element of effective learning environments. *Psychological Test and Assessment Modeling*, 57(2), 262–275.

Torbeyns, J., Verschaffel, L., & Ghesquière, P. (2006). Development of children’s strategy competencies in the number domain from 20 up to 100. *Cognition and Instruction,* 24, 439–465. DOI: 10.1207/s1532690xci2404_2

Törner, G. (1998). Mathematical beliefs and their impact on teaching and learning of mathematics. In Pehkonen, E., & Törner, G. (Eds.), *The State-of-Art in Mathematics-Related Belief Research: Results of the MAVI activities,* 195, 73–94. Helsinki: University of Helsinki.

Verschaffel, L., Greer, B., & De Corte, E. (2007). Whole number concepts and operations. In F. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning*, 557–628. Charlotte: Information Age Publishing Inc.

Xu, L., Liu, R., Star, J.R., Wang, J., Liu, Y., Zhen, R. (2017). Measures of potential flexibility and practical flexibility in equation solving. *Frontiers in Psychology*, 8, 1368. DOI: 10.3389/fpsyg.2017.01368