UNBOUNDED NUMBER LINES REFLECT THE COMPLEXITY OF RATIONAL NUMBER KNOWLEDGE

Sofia Ikonen[1], Jake McMullen[1,2]

[1]Department of Teacher Education, University of Turku
[2]Turku Institute for Learning Analytics, University of Turku

Number line estimation tasks have been found to be beneficial for strengthening rational number knowledge, but there are many open questions regarding the kinds of rational number knowledge that are needed to be proficient in the task. In this study, we aimed to deepen our understanding of a previously less explored unbounded number line estimation task by examining 9th graders' (N=95) qualitatively different answers to unbounded number line estimation tasks, as well as the connections between this and other measures of rational number knowledge. Our results show that students' knowledge appeared fragmented and that performance in the unbounded number line estimation tasks was able to explain variance in students' adaptive rational number knowledge. Implications are discussed.

INTRODUCTION

Rational number knowledge has been identified as instrumental for future mathematical development, and is important for success in many fields of study and in everyday life (Lortie-Forgues et al., 2015). As both students and adults have been found to struggle with rational numbers (Van Dooren et al., 2015), understanding how rational number knowledge develops and how to support that development has long been of interest for researchers and educators. The number line estimation (NLE) task has been used both to measure and support the development of rational number knowledge (Cramer et al., 2019; Van Dooren, 2023). Training with NLE tasks promotes rational number knowledge and NLE skills predict future mathematical success (McMullen et al., 2024; Schneider et al., 2018). However, there is no consensus in the research community on why NLE tasks predict mathematical achievement or what types of knowledge are needed to solve them (Schneider et al., 2018).

In this study, we examine the unbounded NLE task and its connections to other measures of rational number knowledge. The unbounded NLE task (where at least one of the number line's ends is unlabelled) has been theorized to require a wealth of different skills from its solver and proven to promote rational number knowledge in intervention studies (Cramer et al., 2019), including adaptive rational number knowledge (McMullen et al., 2024). Adaptive rational number knowledge refers to "a well-integrated network of knowledge of rational number characteristics and arithmetic relations" (McMullen et al., 2024, p. 2), and is theorized to be a behavioural manifestation of adaptive expertise with rational numbers (McMullen et al., 2020). Adaptive expertise describes being able to apply knowledge in novel contexts with

2026. In H. Krzywacki, M. Hähkiöniemi, E. Haataja, M.S. Hannula & P. Hästö (Eds.). *Proceedings of the 49th Conference of the International Group for the Psychology of Mathematics Education: Research Reports*, pp. 667-674. PME.

understanding, and is contrasted with routine expertise, which refers to the ability to execute familiar tasks efficiently (Hatano, 2003). In the present study, we further consider adaptive rational number knowledge from a *knowledge-in-pieces* theoretical perspective, in which novice knowledge is seen as fragmented and highly context-dependent and learning and expertise are seen as a reorganisation of knowledge structures so that knowledge becomes less context dependent (Levin, 2018). This framework helps us to examine the qualitative differences between adaptive and routine rational number knowledge, and can help to better place adaptive rational number knowledge within theories of learning and knowledge development.

NUMBER LINE ESTIMATION TASKS

There are many different variants of the NLE task (Teppo & van den Heuvel-Panhuizen, 2014). The most commonly known variant is perhaps the bounded NLE task, where participants are asked to place a number on a number line with two labelled end points (typically 0 and 1, 10, 100 etc.) (Schneider et al., 2018). In this study, we used a variant of the unbounded NLE task, where the left endpoint of the number line was labelled with 0 and another, non-integer, point was named somewhere on the number line. Notably, this allowed for the target number and the labelled point on the number line to be of different notations (e.g. a fraction and a decimal). Other variants of the unbounded NLE task might have no end points labelled or not have 0 marked on the number line at all (Cramer et al., 2019; Kim, 2024). Solving unbounded NLE tasks requires coordinating many different skills, including holding together multiple ideas of unit, coordinating numerical and linear units and partitioning and iterating (Cramer et al., 2019; Saxe et al., 2010). This may be one reason why unbounded NLE tasks have been found to support participants' adaptive rational number knowledge (McMullen et al., 2024).

All this suggests that there is reason to pay attention to how to best capture qualitatively different answers and potentially changing performance across different unbounded NLE tasks. Success in the bounded NLE tasks is most often measured with percentage of absolute error (PAE), which is calculated by dividing the absolute distance between the given and correct answer by the length of the number line (Schneider et al., 2018). We propose that measurements that rely purely on the distance between correct and given answer (e.g. PAE) may not be well suited for measuring performance in unbounded NLE tasks. Relative location on the whole number line is not the only relevant measure in unbounded tasks, and more nuanced measurement strategies may be useful (Kim, 2024; Saxe et al., 2010). For example, if 0 and ½ are marked on a number line and the digit to be placed is 0.6, answers at the positions of 0.4 and 0.8 have the same PAE, but the first answer indicates 0.6 would be smaller than ½, while the latter correctly compares 0.6 is bigger than ½, even if the estimations are similarly inaccurate. Thus, the position of the given answer in relation to the other labelled points on the number line may be important to consider separately from pure accuracy measures. In this study, we aim to develop a new scoring framework to capture

potential qualitative differences in students' performance and understanding across unbounded NLE tasks.

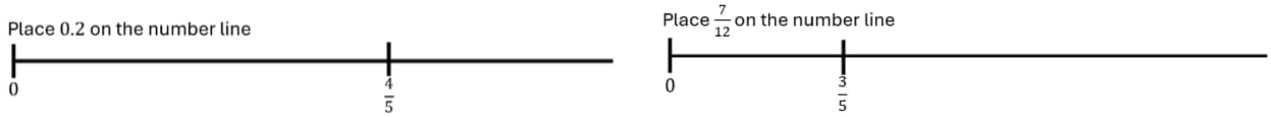
The myriad skills needed to solve unbounded NLE tasks combined with the connection to adaptive rational number knowledge seem to indicate that the task combines many different facets of rational number knowledge and might be a good fit both for understanding mathematical development and for instruction. To date, we are not aware of studies explicitly examining the connections between unbounded NLE tasks using rational numbers with other measures of rational number knowledge. In this study, we aim to better understand unbounded NLE tasks with rational numbers by asking the following questions: (1) How can students' performance on the unbounded NLE tasks be described beyond PAE? and (2) Do students with different performance profiles in the unbounded NLE task differ in their performance in other measures of rational number knowledge?

METHODS

Data

The data for this study was collected in a large, suburban high school in the Southeastern United States. 95 9th graders (50.5% male, age M=14.6) participated in the study. The study was given a positive evaluation from the Ethics Board of the University of Turku and permissions from district, school, parents, and students were acquired. The data was collected during mathematics lessons using a pen and paper questionnaire consisting of different rational number tasks. Three routine rational number knowledge tasks were used: decimal and fraction *arithmetic tasks* (16 items; e.g., $0.75 \times 0.2 =$), *notation conversion tasks* (10 items; e.g., convert 0.375 to fractions) and *ordering tasks* (5 items; e.g., Put the numbers in order from smallest to largest. 0.5; $\frac{1}{4}$; $\frac{5}{100}$; 0.356). The *arithmetic sentence production tasks* measured adaptive rational number knowledge (McMullen et al., 2020); participants were given 90 seconds to produce as many distinct correct arithmetic sentences as possible, using given numbers and equalling the given target number (4 items; e.g., target $\frac{1}{2}$, given numbers $2, \frac{3}{4}$, 0.75, $\frac{1}{4}, 0.25$). Two types of number line estimation tasks were used: *bounded NLE tasks* (8 items, e.g., place 0.42 on a 0–1 number line) and *unbounded NLE tasks* (8 items, see Figure 1).

Figure 1: Unbounded NLE tasks 5 (on the left) and 7 (on the right)



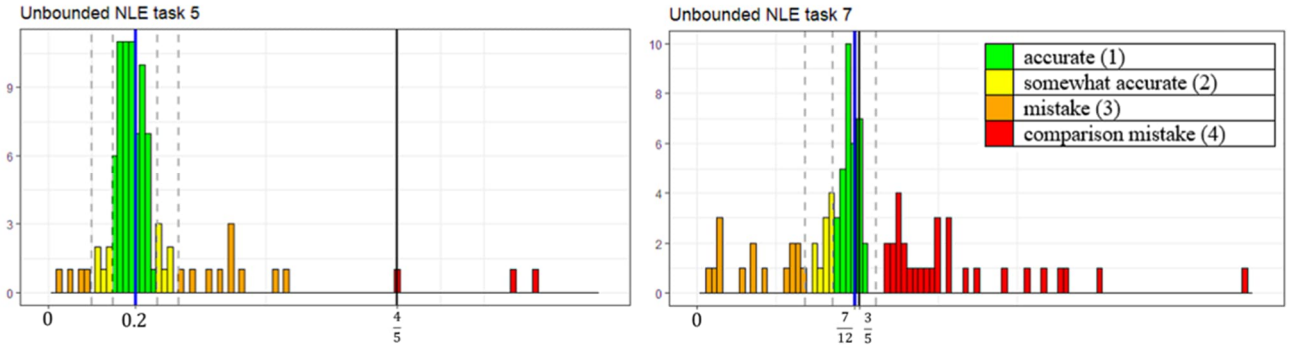
Data analysis

For the analyses, the routine rational number knowledge tasks were scored by calculating the sum of correct answers in each task. In the arithmetic sentence

production tasks, the participants got a point for each correct answer across all four tasks. The bounded NLE tasks were scored with PAE, calculated by dividing the absolute distance between correct and given answer by the length of the number line.

In order to better capture the different components of knowledge needed in the unbounded NLE tasks, we used a combination of qualitative and quantitative methods. We first defined four categories to describe different answers, depending on their interval placement on the number line. These categories took into account the relative location of the response compared to the given number based on the multiplicative relation between given number and target value. The categories were: (1) accurate (correct location ± greatest common divisor of target and marker ÷ 4; Fig. 2 Green)); (2) somewhat accurate (correct ± greatest common divisor of target and marker ÷2; Fig. 2 Yellow), (3) a mistake on the correct side of the given number (Fig. 2 Orange), and (4) a mistake due to a wrong magnitude comparison between marker and target (Fig 2 Red). The exception to these rules was task 7, where limits based on greatest common divisor were difficult to establish (see Fig. 2), and categories 1 and 2 were determined based on the answering pattern and potential measurement error.

Figure 2: Distribution and scoring of answers in unbounded NLE task 5 (on the left) and task 7 (on the right)



Performance-based groups were formed using these categories in order to separate the high and poor performers from the rest of the sample. We placed a participant in the high performance group (n = 29) if they had a maximum of two somewhat accurate (category 2) answers OR one mistake (category 3 or 4) in their answers, with the rest accurate. A participant was placed in the low performance group (n = 16), if they had 4 or more mistakes (category 3 or 4) in their answers. Other participants were placed into the moderate performance group (n = 47).

Analysis of variance tests and post hoc pairwise comparisons using Tukey HSD test were used to determine differences between unbounded NLE performance groups in the other measures of rational number knowledge. Where the assumptions of ANOVA were not met, a Kruskall-Wallis test and Dunn's test were used instead.

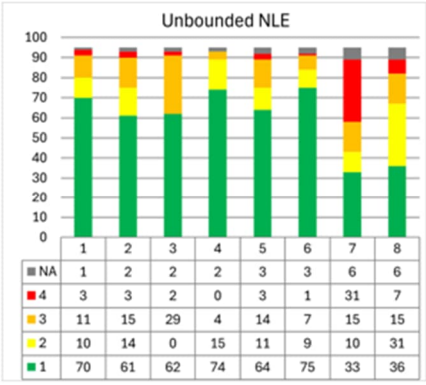
RESULTS

Examination of student performance across the unbounded NLE tasks revealed that the majority of answers were accurate or somewhat accurate (Fig. 3a). Tasks 7 and 8 have

a larger proportion of students making a mistake. Task 7 (place $\frac{7}{12}$ with $\frac{3}{5}$ marked) had the largest proportion of comparison mistakes across all tasks, with 31 students placing the target on the wrong side of the marker (i.e., $\frac{7}{12}$ to the right of $\frac{3}{5}$). Figure 3b reveals that despite the overall accuracy being quite high, on an individual level, students' performance varied markedly from task to task. This pattern was also observed in the moderate and low performance groups. Only 19 students out of the 95 had all responses in the accurate or somewhat accurate category. On the other hand, every student answered accurately to at least two tasks.

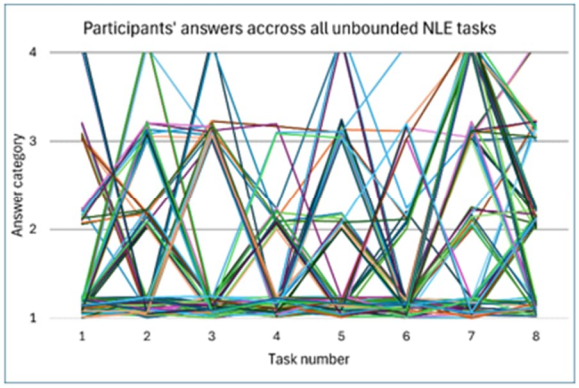
Figure 3 a) Distribution of answers for all unbounded NLE tasks; b) A line graph of all participants' answers to all unbounded NLE tasks

a)



	1	2	3	4	5	6	7	8
NA	1	2	2	2	3	3	6	6
4	3	3	2	0	3	1	31	7
3	11	15	29	4	14	7	15	15
2	10	14	0	15	11	9	10	31
1	70	61	62	74	64	75	33	36

b)



Even in the high performance group (Fig. 4), at least one student made a mistake on four of the tasks. The moderate performance group was allowed more mistakes, most of which were category 3 mistakes. Outside of task 7, the mistakes and inaccuracies were distributed somewhat evenly between the tasks in this group. The low performance group's mistakes had more comparison errors compared to the moderate performance group, and notably there were no accurate or somewhat accurate (category 1 or 2) answers in task 7. Examining the within-groups answers revealed no patterns that could reliably explain group differences, as answering patterns within groups remained just as unpredictable as at the full-sample level.

Figure 4: Stacked bar charts for each unbounded NLE performance group's answers to each unbounded NLE task

One-way ANOVAs were performed to answer research question 2. Levene's test of equality of variances indicated that equal variances between groups could not be assumed ($F(2,89)=4.2$, $p=.02$ and $F(2,89)=13.5$, $p<.001$, respectively) for ordering tasks and bounded NLE tasks, so a more robust Kruskall-Wallis test was used. Performance groups differed for all measures of rational number knowledge (Table 1). A post hoc Tukey HSD test indicated that the mean of number of produced arithmetic sentences of the high performance group was significantly higher than that of the moderate ($p=.009$) and low ($p<.001$) performance groups, and that of the moderate performance group was significantly higher than the low performance group's ($p=.003$). All other post-hoc comparisons followed the pattern of the high and moderate performance groups not having statistically significant differences (ps > .05), but both having higher performance than the low performance group (ps < .05).

Table 1: Summary statistics of each unbounded NLE performance group (H=high; M=moderate; L=low) in the other measures, as well as results of analysis of variance and post hoc tests. *** $p<.001$; [1]Tukey HSD; [2] Dunn test

	Group means (SD)			ANOVA	post-hoc
	H (n=29)	M (n=47)	L (n=16)		
arithmetic (max. 16)	9.4 (3.6)	7.8 (3.6)	4.5 (2.8)	$F(2,89) = 10.2^{***}$	H=M>L[1]
notation conversion (max. 10)	7.6 (1.9)	6.7 (2.2)	3.7 (2.2)	$F(2,89) = 18.6^{***}$	H=M>L[1]
ordering (max. 5)	4.0 (0.9)	3.5 (1.4)	2.1 (1.3)	$\chi^2(2, N = 92) = 18.1^{***}$	H=M>L[2]
arithmetic sentence production	17.8 (6.0)	13.8 (5.5)	8.6 (5.4)	$F(2,89) = 13.8^{***}$	H>M>L[1]
bounded NLE	.045 (.025)	.061 (.032)	.108 (.066)	$\chi^2(2, N = 92) = 15.7^{***}$	H=M>L[2]

DISCUSSION

Overall, most students got an accurate or somewhat accurate score on more than half of the eight unbounded NLE tasks. However, the overall rational number knowledge reflected in the unbounded NLE tasks seems to be fairly fragmented; even within groups, individual students' response patterns appeared haphazard and high performing students made mistakes. This might be evidence of the piecemeal nature of students' rational number knowledge: instead of an overarching skill in determining the relationship between the two numbers in a given task, students' knowledge seems

to be more number (context) dependent, with success in one task not necessarily translating into success in another, which would be in line with the knowledge in pieces framework's view of novice knowledge (Levin, 2018).

The idea that well-integrated knowledge is central for success in the unbounded NLE task is further supported by the result that adaptive rational number knowledge was the only measure where post hoc tests revealed differences between the high and moderate performance groups. This seems to suggest that adequate skill in routine measures of rational number knowledge (e.g., fraction to decimal conversion) was necessary but not sufficient to perform well in the unbounded NLE tasks. Subsequently, differences between the high and moderate performance groups might be attributed to qualitative differences in the students' knowledge, namely more integrated knowledge. These results are in line with previous research indicating that the unbounded NLE tasks require a wealth of different knowledge and therefore may require more well-integrated knowledge for consistent success (Cramer et al., 2019; McMullen et al., 2024; Saxe et al., 2010).

In this paper we introduced a new way of measuring performance on unbounded NLE tasks. The scoring method we developed in this study takes into account some of the characteristics of unbounded NLE that the methods used for bounded NLE do not, all the while allowing for scoring each participants' answers based on numerical guidelines. The results of this study indicate that the new scoring method was able to differentiate participants from one another, and performance groups formed based on the scoring differed from one another in other measures of rational number knowledge. However, methodically comparing the new method to PAE and verifying the method based on actual strategies used were outside of the scope of the present study.

The results of this study support previous research's claim that the unbounded NLE task is complex and not yet completely understood. We also found a previously undiscovered connection between unbounded NLE tasks and adaptive expertise with rational numbers. Our results indicate that the unbounded NLE task could be used to examine and develop students' adaptive expertise with rational numbers, but it is important to note that the scope of this research is limited by the relatively small number of unbounded NLE tasks, and more research is needed to understand how students' knowledge is reflected in solving the tasks.

Acknowledgements

The present study is a part of the EDUCA Flagship and SCAFFLEX project funded by the Research Council of Finland (#358924, #358947, #363159) and supported by the Jacobs Foundation.

References

Cramer, K., Monson, D., Ahrendt, S., Wyberg, T., Pettis, C., & Fagerlund, C. (2019). Reconstructing the unit on the number line: Tasks to extend fourth graders' fraction

understandings. *Investigations in Mathematics Learning*, *11*(3), 180–194. https://doi.org/10.1080/19477503.2018.1434594

Hatano, G. (2003). Foreword. In A. Baroody, & A. Dowker (Eds.), *The Development of Arithmetic Concepts and Skills: Constructing Adaptive Expertise* (pp. xi–xiii). Routledge.

Kim, J. (2024). Profiles of young students' understanding of fractions on number lines. *Eurasia Journal of Mathematics, Science and Technology Education*, *20*(5), Article em2444. https://doi.org/10.29333/ejmste/14469

Levin, M. (2018). Conceptual and Procedural Knowledge During Strategy Construction: A Complex Knowledge Systems Perspective. *Cognition and Instruction*, *36*(3), 247–278. https://doi.org/10.1080/07370008.2018.1464003

Lortie-Forgues, H., Tian, J., & Siegler, R. S. (2015). Why is learning fraction and decimal arithmetic so difficult? *Developmental Review*, *38*, 201–221. https://doi.org/10.1016/j.dr.2015.07.008

McMullen, J., Hannula-Sormunen, M. M., Lehtinen, E., & Siegler, R. S. (2020). Distinguishing adaptive from routine expertise with rational number arithmetic. *Learning and Instruction*, *68*, Article 101347. https://doi.org/10.1016/j.learninstruc.2020.101347

McMullen, J., Koskinen, A., Kärki, T., Lindstedt, A., Määttä, S., Halme, H., Lehtinen, E., Hannula-Sormunen, M. M., & Kiili, K. (2024). A game-based approach to promoting adaptive rational number knowledge. *Mathematical Thinking and Learning*, *26*(4), 411–427. https://doi.org/10.1080/10986065.2023.2177818

Saxe, G. B., Earnest, D., Sitabkhan, Y., Haldar, L. C., Lewis, K. E., & Zheng, Y. (2010). Supporting Generative Thinking About the Integer Number Line in Elementary Mathematics. *Cognition and Instruction*, *28*(4), 433–474. https://doi.org/10.1080/07370008.2010.511569

Schneider, M., Merz, S., Stricker, J., De Smedt, B., Torbeyns, J., Verschaffel, L., & Luwel, K. (2018). Associations of Number Line Estimation With Mathematical Competence: A Meta-analysis. *Child Development*, *89*(5), 1467–1484. https://doi.org/10.1111/cdev.13068

Teppo, A., & Heuvel-Panhuizen, M. (2014). Visual representations as objects of analysis: The number line as an example. *ZDM*, *46*(1), 45–58. https://doi.org/10.1007/s11858-013-0518-2

Van Dooren, W. (2023). Strategy use in number line estimations of fractions—An explorative study in search for adaptive expertise. In M. Ayalon, B. Koichu, R. Leikin, L. Rubel, & M. Tabach (Eds.), *Proc. of the 46th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 4, pp. 299–306). PME.

Van Dooren, W., Lehtinen, E., & Verschaffel, L. (2015). Unraveling the gap between natural and rational numbers. *Learning and Instruction*, *37*, 1–4. https://doi.org/10.1016/j.learninstruc.2015.01.001