INDIVIDUAL DIFFERENCES IN FRACTIONS' CONCEPTUAL AND PROCEDURAL KNOWLEDGE: WHAT ABOUT OLDER STUDENTS?

Maria Bempeni*, Stavroula Poulopoulou**, Ioanna Tsiplaki*, and Xenia Vamvakoussi*

*University of Ioannina, **Athens University of Economics and Business

We constructed and calibrated an instrument targeting conceptual and procedural fraction knowledge. We used this instrument in a quantitative study with 126 secondary students (7th and 9th graders), testing the hypothesis that there are individual differences in the way students combine the two types of knowledge. Cluster analysis revealed four distinct student profiles: Students who were either stronger or weaker than expected with respect to both types of knowledge; students who were stronger with respect to conceptual knowledge; and students who were stronger with respect to procedural knowledge. These findings support the individual differences hypothesis.

THEORETICAL BACKGROUND

Procedural knowledge typically refers to the ability to execute action sequences to solve problems and is usually tied to specific problem types, whereas conceptual knowledge is defined as the knowledge of concepts and principles pertaining to a domain (Rittle-Johnson, Siegler, & Alibali, 2001; but see also Star (2005) for a plea to reconsider how procedural knowledge is conceptualized). Research in the area highlights the fact that conceptual and procedural mathematical knowledge (hereafter, CKn and PKn) are equally important for mathematical competence (Rittle-Johnson & Schneider, 2015), yet mathematics education wavers between giving precedence to one or the other type of knowledge (Moss & Case, 1999; Star, 2005). As a consequence, the issue of procedural and conceptual knowledge in mathematics learning is not only theoretically interesting, but also educationally relevant.

There has been a lot of discussion regarding which type of knowledge develops first. Procedures-first theories assume that children first learn procedures for solving problems in a domain and then derive domain CKn from their experience solving problems. Concept-first theories support that students initially acquire CKn and then build PKn through repeated practice (Rittle-Johnson et al., 2001). Noting that, regardless of which type of knowledge comes first, the relation between CKn and PKn is typically bi-directional, Rittle-Johnson and colleagues (2001) argued for an iterative model, according to which the two types of knowledge develop in a hand-over-hand process and gains in one type of knowledge lead to improvements in the other, which in turn increases the first type of knowledge.

2018. In E. Bergqvist, M. Österholm, C. Granberg, & L. Sumpter (Eds.). *Proceedings of the 42nd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 147-154). Umeå, Sweden: PME.

However, the relations between CKn and PKn remain a complex issue: For example, increases in one type of knowledge do not always result in equal amount of increase in the other; moreover, it appears that there are individual differences in the way students combine the two types of knowledge (Rittle-Johnson & Schneider, 2015). Hallett and colleagues (Hallett, Bryant, & Nunes, 2010; Hallett, Nunes, Bryant and Thrope, 2012) investigated such individual differences in students' fraction knowledge. They assessed CKn and PKn of students at Grade 4 and 5 (2010) as well as at Grade 6 and 8 (2012) and identified groups of students who had strong (or weak) CKn as well as PKn. However, they also consistently traced two substantial groups of students who demonstrated relative strength with one form of knowledge and weakness with the other, with differences between the two types of knowledge becoming less salient with age.

In an in-depth qualitative study, we found, similarly to Hallett and colleagues (2010, 2012), individual differences in the extent to which 9th graders rely on CKn and PKn in the area of fractions (Bempeni & Vamvakoussi, 2015). More specifically, we found students who displayed flawless procedural performance but failed even in the simplest of tasks that required conceptual understanding of fractions, and vice versa. Such findings are theoretically interesting because they challenge the view that all children follow a uniform sequence in gaining the two types of knowledge (see also Canobi, 2004) and illustrate the possibility that CKn and PKn may not develop in a hand-over-hand manner, putting a challenge to the Rittle-Johnson et al.'s iterative model (2001). Moreover, tracing salient individual differences at grade 9 could indicate that individual differences may persist, despite the general tendency to diminish overtime (Hallett et al., 2010; 2012).

An issue that needs to be addressed in this research area is the fact that very little attention has been given to measurement validity. As Rittle-Johnson & Schneider (2015, p. 1128) noted:

> However, before more progress can be made in understanding the relations between conceptual and procedural knowledge, we must pay more attention to the *validity* of measures of conceptual and procedural knowledge. Currently, no standardized approaches for assessing conceptual and procedural knowledge with proven validity, reliability, and objectivity have been developed.

To further investigate this issue, we constructed and calibrated a new instrument measuring CKn and PKn of fractions. We administered this instrument to secondary students testing the hypothesis that there are individual differences in the way students combine CKn and PKn that remain salient at the secondary level.

METHOD

Research instrument

The research instrument—in its final form after the evaluation (see Instrument Evaluation below)—comprised 26 fraction tasks grouped in two categories, procedural

(12) and conceptual (14) tasks. The procedural tasks were paper-and-pencil tasks requiring knowledge of procedures taught at school (e.g., to execute fractions operations, to find an equivalent fraction, to cross-multiply, to simplify a complex fraction, and to compare dissimilar fractions).

The conceptual tasks were based on our materials from previous studies (Bempeni & Vamvakoussi, 2015) as well as on other instruments assessing conceptual understanding of fractions (e.g., Van Hoof, Verschaffel, & Van Dooren, 2015) and targeted many important aspects of fraction CKn. For example, students were asked to interpret and evaluate fraction representations with area models as well as the number line; to mentally compare fractions; to estimate the outcome of fraction operations; and to select appropriate fraction operations to solve problems.

The conceptual tasks were posed in the form of multiple choice questions in order to ensure that students would not use paper-and-pencil. This is because some conceptual tasks could be tackled with procedural strategies (e.g., the comparison of fractions), in which case students' PKn rather than their CKn would be assessed (see also Rittle-Johnson & Schneider, 2015).

Participants

The participants of the study were 126 Greek students: 66 seventh graders and 60 ninth graders.

Procedure

The same questionnaire was issued in two versions (A and B), varying the order of presentation of the conceptual tasks so as to prevent students from cheating. The students had fifty minutes to solve the fractions tasks.

DATA ANALYSIS AND RESULTS

The responses of the tasks were coded as correct or wrong. For the data analysis we used the Statistical Package for the Social Sciences (SPSS) and the R Project for Statistical Computing.

Instrument Evaluation

We conducted a clinical pilot with 61 students in order to evaluate the reliability and the validity of our instrument. The instrument, in its initial form, included 39 fraction tasks (12 procedural and 27 conceptual tasks).

The instrument showed strong *face validity* given that all tasks were assessed as clear, reasonable and accurate by 6 mathematics education experts. The experts were also asked to rate the relevance of each item to the aim of the instrument, on a 4-point scale. The calculation of the *content validity* index (CVI=1>0.83) for each item confirmed the high consistency between experts (Polit, Beck, & Owen, 2007). Moreover, multitrait analysis illustrated that all the items of the PKn scale showed *convergent validity* and *divergent validity* by demonstrating high correlation with the procedural scale and low correlation with the conceptual scale respectively. More specifically, the value of correlation for all procedural items was above 0.7. Eight items of the CKn scale showed low correlation with conceptual tasks or higher than expected correlation with PKn tasks, possibly due to their great diversity (see also Hallett et al., 2012). We decided to exclude these items from subsequent analyses.

To establish whether the items on this questionnaire all reliably measure the same construct we calculated Cronbach's alpha. The instrument showed a high degree of *internal consistency* (0.921 and 0.731 for the PKn and CKn measure, respectively). We also calculated Cronbach's alpha for CKn and PKn scales in case we removed a particular task. In all cases the value of Cronbach's alpha was less than 0.921 and 0.731 for the PKn and CKn scales respectively and as a result there was no reason to delete any of the items (Cronbach, 1951). We also assessed the *external consistency* of the instrument over a period of three weeks with a test-retest method. The value of intra-class correlation coefficient was high for all the PKn tasks ($r>0.8$). However, this was not the case for all the CKn tasks. Five of them presented intra-class correlation coefficient below 0.5. We thus decided to remove them from the final form of our instrument (Ware & Gandek, 1998).

Main study

Our data were analyzed using cluster analysis. Following Hallett et al. (2010), we used the residualized scores of PKn and CKn in our analysis, the raw scores being the percentages of correct answers out of the total of answered questions. This is because the two scales (CKn and PKn) are expected to be correlated, and the use of residualized scores provides a way of measuring CKn and PKn that excludes this common part of variation from both scales (Cohen, Cohen, West, & Aiken, 2003). The residualized scores were obtained using linear regression. More specifically, residuals for CKn were obtained by regressing conceptual knowledge against procedural knowledge, while the residuals for PKn were obtained by regressing PKn against CKn. The residualized scores represent relative, rather than absolute, strength with respect to PKn and CKn, in the sense that a positive residual, for instance in PKn, means that a person's PKn is stronger than expected given their CKn.

In order to identify different profiles of individual differences a cluster analysis was performed based on the two residualized scales using the k-means method and Euclidean distance as a distance measure. In the literature, a wide variety of indices have been proposed to find the optimal number of clusters in a partitioning of a data set during the clustering process. In order to determine the optimal number of clusters at our dataset we used the R package NbClust which provides 26 indices (Brock, Pihur, Datta, & Datta, 2008). The majority of these indices suggested a four cluster solution. The mean (and standard deviation) of the conceptual and procedural scores for the students of the four clusters are presented in Table 1. Table 1 also presents the mean and standard deviation for raw and residualized scores by cluster. The first cluster *Stronger than expected in CKn* (N=21, 16.7%) is characterized by positive CKn residuals and negative PKn residuals which means that students in this cluster performed better than expected in CKn tasks, given their performance in PKn tasks.

	Cluster 1		Cluster 2		Cluster 3		Cluster 4		
	Stronger than expected in CKn		*Stronger than expected in PKn and CKn*		*Stronger than expected in PKn*		*Weaker than expected in PKn and CKn*		
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
Procedural Residual	-1,4	0,4	0,4	0,2	0,8	0,4	-0,7	0,4	<.0001
Conceptual Residual	0,8	0,5	1,3	0,8	-0,7	0,4	-0,3	0,3	<.0001
Procedural Score	10,3	12,8	85,9	10,4	70,1	16,6	19,6	14,8	<.0001
Conceptual Score	30,9	8,5	54,5	14,2	17,5	8,3	13,3	6,3	<.0001

Table 1: Mean and standard deviation for raw and residualized scores by cluster

The second cluster *Stronger than expected in PKn and CKn* (N=22, 17.5%) and fourth cluster *Weaker than expected in PKn and CKn* (N=31, 24.6%) cluster comprised students with either good or poor performance, respectively, in both measures. The third cluster *Stronger than expected in PKn* (N=52, 41.3%) is characterized by negative CKn residuals and positive PKn residuals which means that students in this cluster performed better than expected in PKn tasks, given their performance in CKn tasks.

We present the cases of two students illustrating extreme individual differences regarding PKn and CKn. The student of the *Stronger than expected in PKn* profile achieved 100% score in procedural tasks but only 14.29% in conceptual tasks, failing even in the simplest ones (e.g.:area model, Figure 2).

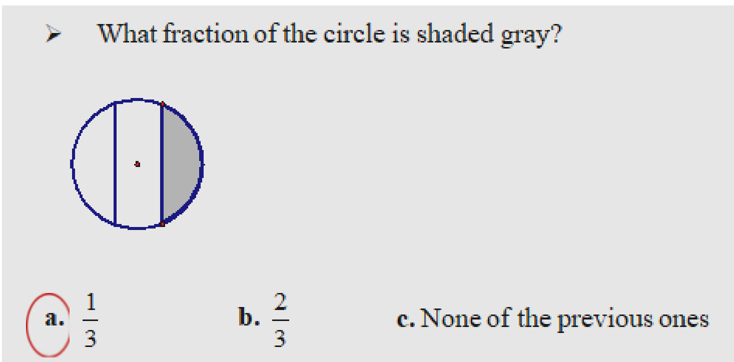

Figure 1: An extreme case in the *Stronger in PKn* cluster

On the other hand, the student of the CKn profile failed in all PKn tasks and responded correctly in 50% of the CKn tasks. Figure 2 illustrates the fact that this particular student was able to estimate a sum of fractions, despite the fact that she had failed to perform fraction addition.

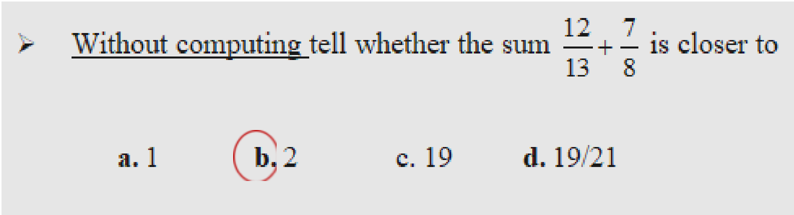

Figure 2: An extreme case in the Stronger in CKn cluster

The chi-square test of independence showed that there were no significant differences in the distribution of the two age groups among the clusters ($\chi^2 = 0.31$, $df = 3$, $p\text{-}value = 0.96$). The distribution of cluster membership across grade is presented in Table 2. The mean raw scores of the two age groups in CKn & PKn tasks are presented in Figure 3.

Grade	Cluster 1 *Stronger than expected in CKn*	Cluster 2 *Stronger than expected in PKn and CKn*	Cluster 3 *Stronger than expected in PKn*	Cluster 4 *Weaker than expected in PKn and CKn*
Seven	10	12	27	17
Nine	11	10	25	14
Total	21	22	52	31

Table 2: Distribution of clusters across grade

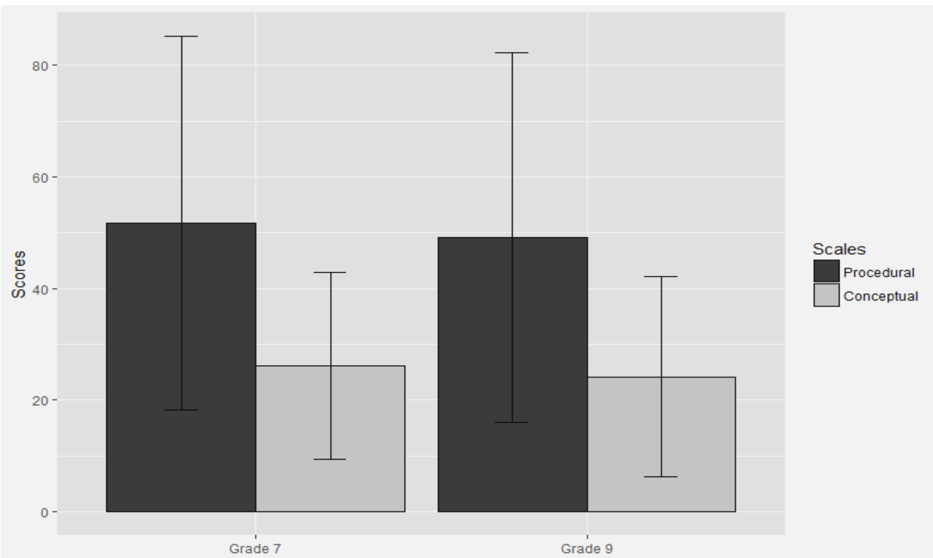

Figure 3: Mean raw scores of the two age groups

CONCLUSIONS-DISCUSSION

In response to the need for valid and reliable measures of conceptual and procedural mathematical knowledge, we constructed and calibrated an instrument targeting conceptual and procedural fraction knowledge. The instrument we developed demonstrated good indicators of validity and reliability and thus can be characterized as an efficient instrument for fraction CKn and PKn.

This instrument was used in a study testing the hypothesis that there are individual differences in the way that students combine CKn and PKn of fractions that remain salient at the secondary level. The results supported this hypothesis. We point out that 43.2% of our sample belonged to the *Stronger than expected in CKn* and *Stronger than expected in in PKn* clusters. This hypothesis is further corroborated by the fact that the student profiles obtained in our study were very similar to Hallett and colleagues' (2010, 2012), despite the fact that a different instrument was used, in a different population (Greek students), and for older students (9^{th} graders). In contrast to Hallett et al. (2012), our findings showed that individual differences in CKn and PKn do not necessarily diminish over time and may even be extreme, as indicated by the examples presented above.

Our findings also indicated that only few students adequately combine CKn and PKn. Given that developing both CKn and PKn is critical for mathematical development, it is important that teaching strategies and techniques that support both types of knowledge are used in instruction (see Rittle-Johnson & Schneider, 2015, for a review). We note that the third cluster *Stronger than expected in PKn* comprised the greater part of our sample, suggesting that more attention should be paid to conceptual knowledge during instruction. Moreover, despite the fact that we expected school experience to lead to improvements in students' fraction knowledge, we did not find any significant differences in the distribution of the two age groups across the clusters.

Further, the validation of such individual differences showcases the importance of differential instruction, based on the students' needs. Students who belong to different profiles could gain from different instructional approaches that make good use of the CKn and PKn that students have (Gilmore & Bryant, 2006).

Finally, the instrument we constructed and calibrated can be used for the assessment of students' understanding of fractions and of the lack in CKn and PKn so as to differentiate appropriately mathematics instruction.

Acknowledgements

This project is implemented through the Operational Program "Human Resources Development, Education and Lifelong Learning" and is co-financed by the European Union (European Social Fund) and Greek national funds.

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο

Επιχειρησιακό Πρόγραμμα
Ανάπτυξη Ανθρώπινου Δυναμικού, Εκπαίδευση και Διά Βίου Μάθηση
Ειδική Υπηρεσία Διαχείρισης
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

ΕΣΠΑ 2014-2020
ανάπτυξη - εργασία - αλληλεγγύη

References

Bempeni, M. & Vamvakoussi, X. (2015). Individual differences in students' knowing and learning about fractions: Evidence from an in-depth qualitative study. *Frontline Learning Research, 3,* 17-34.

Brock, G., Pihur, V., Datta, S., & Datta, S. (2008). clValid: An R package for cluster validation. *Journal of Statistical Software, 25*(4) (p.1-22).

Canobi, K. H. (2004). Individual differences in children's addition and subtraction knowledge. *Cognitive Development, 19,* 81-93.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). Applied multiple regression/correlation analysis for the behavioral sciences (3rd ed.). Mahwah, NJ: Erlbaum.

Cronbach L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 6,* 297–334.

Gilmore, C. K., & Bryant, P. (2006). Individual differences in children's understanding of inversion and arithmetical skill. *British, Journal of Educational Psychology, 76,* 309–331.

Hallett, D., Nunes, T., & Bryant, P. (2010). Individual differences in conceptual and procedural knowledge when learning fractions. *Journal of Educational Psychology, 102,* 395–406.

Hallett, D., Nunes, T., Bryant, P., & Thorpe, C. M. (2012). Individual differences in conceptual and procedural fraction understanding: The role of abilities and school experience. *Journal of Experimental Child Psychology, 113,* 469-486.

Moss, J., & Case, R. (1999). Developing children's understanding of the rational numbers: A new model and an experimental curriculum, *Journal for Research in Mathematics Education, 30,* 122-147.

Polit, D., Beck, C., & Owen, S. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Res Nurs Health, 30,* 459-467.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology, 93,* 346-362.

Rittle-Johnson, & B., Schneider, M. (2015). Developing conceptual and procedural knowledge of mathematics. In R. Kadosh & A. Dowker (Eds.), *Oxford Handbook of Numerical Cognition* (pp.1118-1134). Oxford: Oxford University Press.

Star, J.R. (2005). Reconceptualizing procedural knowledge. *Journal for research in mathematics education, 36,* 404-411.

Ware J. & Gandek B. (1998). Methods for testing data quality, scaling assumptions and reliability: The IQOLA project approach. *J Clin Epidemiol., 51,* 945–952.

Van Hoof, J., Verschaffel, L., & Van Dooren, W. (2015). Inappropriately applying natural number properties in rational number tasks: Characterizing the development of the natural number bias through primary and secondary education. *Educational Studies in Mathematics, 89,* 39–56.