Theresa Kruse, Ulrich Heid, and Barbara Schmidt-Thieme

MATHEMATICS STUDENTS AS LEXICOGRAPHERS

Learning Domain Concepts and Their Relations by Designing Dictionary Articles and Concept Maps

Abstract We present a study which was carried out with teacher students of mathematics. They were asked to create either dictionary articles or concept maps for terms from an introductory lecture in their first semester. Based on the students' submissions, we investigate whether there is a difference in the learning outcomes between the two tasks and also whether the technical means used to solve these tasks influence the students' engagement in the tasks, i.e., whether they chose digital tools or handwriting to complete them. The analysis presented here is based on a first annotated subset of our data and provides preliminary results on our research questions. We show that digital tools seem to be more appropriate to motivate a deeper exploration of the domain. In addition, our analysis suggests that dictionary articles and concept maps motivate different cognitive approaches to the domain, depending on whether the focus is on the concepts themselves or on the possible relations between them.

Keywords learning; mathematics; concept maps; dictionary articles

1. Introduction

In introductory lectures in universities, mathematics tends to be presented as a sequence of definitions, theorems, and proofs, while in school it is more often presented as a collection of algorithms. Our goal is to make first-year students aware of the fact that mathematics is actually a web of concepts that are connected by different kinds of relations. We hypothesize that this can be achieved if students have to describe concepts and the relations between them in the way in which they are presented in dictionaries and concept maps. Thus, we gave the students the task of writing dictionary articles and creating concept maps.

In this paper, we focus on three research questions:

> (RQ1) Is there a relation between mathematical knowledge and engagement[1] in solving dictionary or concept maps tasks, respectively? And if so, how strong is it?
>
> (RQ2) Are there differences in the engagement between dictionary and concept map tasks? And if so, what are they?
>
> (RQ3) Does the engagement in dictionary and concept map tasks depend on the mean used to solve them? And if so, to what extent?

[1] By engagement we mean here the intensity with which the students have worked on the task, measured in the number of mentioned concepts and relations.

This paper is part of the publication: Despot, K. Š., Ostroški Anić, A., & Brač, I. (Eds.). (2024). *Lexicography and Semantics. Proceedings of the XXI EURALEX International Congress.* Institute for the Croatian Language.

We present preliminary answers to these questions based on a subset of the data we collected: We look at the first tasks handed in by the students. We present the setting in which we conducted our study in Section 2, while the actual analysis is given in Section 3. Section 4 summarizes the results and our learnings. Limitations and a final discussion are given in Section 5.

2. Experimental Setting

The participants of our study are students in the first year of their studies to become mathematics teachers. The course was carried out in the winter term of 2023/24. At the beginning of the semester, 167 students enrolled in the course, out of which 44 were assigned to create concept maps, 50 to create dictionary articles and 73 did other terminology-related tasks. The allocation of the students to the three groups was not randomized but depended on the time slot they chose for their tutorial session. Thus, external factors such as time slots occupied by courses in their second field of study may have influenced the distribution. This also explains the different group sizes.

Each week, the students were given a choice task and a compulsory task. The choice tasks were the same for all students and comparable to classical mathematical tasks at university, i.e., calculating or proving things. In the following evaluation, we use the choice task for estimating the general mathematical abilities. The compulsory tasks differed between the students. The first group always had the task of writing dictionary articles for two given concepts, the second group always had the task of creating a concept map containing the two same concepts, and the third group was a control group that was given either the dictionary task, the concept map task, or another terminology task. There were ten of these tasks over the semester. On average, 110 people submitted their solutions each week. This number is lower than the 167 persons who originally signed up for the course because of dropouts and course repeaters. Over the semester, we collected a total of 470 submissions for the dictionary article tasks, 329 submissions for the concept maps, and 303 submissions for the other tasks. At the beginning of the semester, the students were given a short introduction to the production of dictionary entries and concept maps, covering both, structural and infrastructural aspects, such as recommended software. They were also informed about the fact that we conduct a study accompanying the course.

3. Analysis of Students' Submissions

In this section, we present a first analysis of the students' submissions from which we generate hypotheses for answering the research questions which have to be evaluated on the whole data set.

4. Dictionary Articles

The purpose of the dictionary articles can be described in the framework of the function theory (Fuertes-Olivera & Tarp, 2014). In our case, the users are the same group as the lexicographers as they are writing the dictionary entries for themselves,

This paper is part of the publication: Despot, K. Š., Ostroški Anić, A., & Brač, I. (Eds.). (2024). *Lexicography and Semantics. Proceedings of the XXI EURALEX International Congress.* Institute for the Croatian Language.

for example for using them as a learning tool at a later point. Further users are the lecturers as they evaluate the results, but as this is part of the meta-setting, we do not count the lecturers as users in the narrower sense. The students are writing the entries in (one of) their first language(s). Concerning the lexicographic aspects, they are lay people, and they are semi-experts in the domain of the entries (Bowker, 2003). The usage situation is a cognitive one as its main purpose is to gain or deepen knowledge about the domain. The situations could be both, sporadic and systematic, as there could be sporadic situations where a user wants to look up a certain meaning, while the overall interest is a systematic one, i.e., learning and understanding the concepts of the given domain.

The following data is based on the first dictionary task in which the students were asked to create articles for the terms *Wertemenge* (i.e., *range of a function*) and *Definitionsmenge* (i.e., *domain of a function*). 40 persons handed in their solutions.

First, we look at the items they use in their article structures and give the percentage of students who used a given item in their article structure in parentheses: definition (100%), synonyms (82.5%), hypernyms (82.5%), hyponyms (77.5%), associated concepts (77.5%), collocations (70%), Part of Speech labels (PoS) (5%). Two further item types, namely 'further information' and 'variant' are used by one person each. Besides PoS, all the information provided belongs to the comment on semantics and is mostly identical to typical items in specialized dictionaries (Humbley, 2017).

We also annotate the technical mean used by the students to create the articles and find five categories: editor (42.5%), XML (27.5%), tablet (20%), hand-written (5%), table (5%). With the category "editor" we refer to dictionary entries written with a text editing program. For highlighting the structural indicators, different strategies like bold typeface are used. In the table category, the entries are provided in a table. The XML entries have been created with Lexonomy (Měchura, 2017) and exported. The category 'tablet' comprises entries created using a tablet, which is a variant of the use of handwriting but provides a wider range of possibilities for a more flexible editing due to its digital nature. This flexibility distinguishes them from the submissions in the category hand-written which are scans or photos of hand-written articles on paper.

All students use a panel view for presenting their data (Koplenig & Müller-Spitzer, 2014). Single students also create a kind of access structure, e.g., by an index, but this was not part of the task they got.

5. Concept Maps

We analogously evaluate the first task for the group who created the concept maps for the same two terms, *Wertemenge* and *Definitionsmenge*. It was handed in by only 25 persons. For the annotation of the categories, we use a deductive approach based on a category system for relations between mathematical terms by Kruse et al. (2023). The most common categories in the task we analyze are (with the percentage of

This paper is part of the publication: Despot, K. Š., Ostroški Anić, A., & Brač, I. (Eds.). (2024). *Lexicography and Semantics. Proceedings of the XXI EURALEX International Congress.* Institute for the Croatian Language.

students who used them given in parentheses): synonymy (92%), hypo-/hyperonymy (84%), part-of-relation (80%), mapping (32%), causality (32%), analogy (8%), property (4%), medium (4%). In addition, there are connection lines in the maps without a named relation (44%) and others that could not be mapped to one of the categories of the scheme but have been only used by single students (24%).

Analogously to the dictionary task, we also evaluate the means the concept maps were created with. We find the following categories analogously to the dictionary entries: hand-written (24%), with a tablet (44%), or with a digital tool for creating concept maps (32%).

6. Results of Quantitative Evaluations and Learnings

For the quantitative analysis of the students' submissions, we carry out t-tests for comparing different subgroups and we calculate correlation coefficients for finding relations between different aspects. As a reference, we use the points achieved in the choice task which was handed in with the dictionary or concept map tasks, respectively ('points'). Furthermore, we consider the number of different items in the article structure for the dictionary task which we regard as equivalent to the number of different relations in the concept map task ('relations'). The third number we consider is the total number of items attributed to the structural indicators ('items').

First, we look at correlations between the different values. For the concept maps, we find no significant pairwise correlations (see Table 1). The reason for that could be that we only have a small amount of data in that first task, with 25 participants. However, the correlations for the dictionary task are significant and we find in all three cases positive correlations with medium to large effect (Cohen, 1988). It would be interesting to see if the numbers for the concept maps give similar results once calculated for the bigger data set. For now, we can conclude that students who acquire more points in other mathematical tasks tend to create bigger dictionary articles. Nevertheless, this correlation might also be caused by other factors not considered in our survey.

Table 1: Results of Pearson correlations between items, relations and points achieved

	Dictionary articles (N=40)	**Concept maps (N=25)**
Items and relations	0.8371 ($p<0.001$)	0.3492 ($p=0.0871$)
Items and points	0.4768 ($p<0.01$)	-0.1188 ($p=0.5716$)
Relations and points	0.5989 ($p<0.001$)	-0.3629 ($p=0.0746$)

Thus, we investigate if there are significant differences between both groups. The results in Table 2 show that there is no significant difference in the achieved points (out of 20) between the two groups. However, in the dictionary articles, significantly more different relations are used than in the concept maps. The reason for that could

This paper is part of the publication: Despot, K. Š., Ostroški Anić, A., & Brač, I. (Eds.). (2024). *Lexicography and Semantics. Proceedings of the XXI EURALEX International Congress.* Institute for the Croatian Language.

be that both formats favor different approaches in the creation, as someone who is writing a dictionary article might be more focused on adding more items to the article structure, while someone creating a concept map is more focused on adding more concepts rather than finding more different relations. This could be also reflected by the fact that there is no significant difference in the total number of items given between both groups.

Table 2: Arithmetic means and t-test results

	Dictionary articles (N=40)	**Concept maps (N=25)**	**Significance**
Points	14.63	15.98	n.s.
Relations	5.20	4.04	$p<0.01$
Items	19.5	17.2	n.s.

The third aspect we evaluate is the mean which the students use to create their dictionary articles or concept maps, respectively. Therefore, we divide the categories into two groups, namely digital and hand-written. The idea behind this is mainly that students who are typing are more flexible in their editing than when doing it by hand where they are more bound to an established structure. Nevertheless, we have to keep in mind that our analyses are only based on the final submissions and we do not have access to the students' work processes where different draft versions might have been created.

Concerning the amount of points reached in the choice tasks, we find no significant differences in the preference for one or the other mean. Significant differences are only found for the dictionary entries where more relations and items are added in the digital tools, while the results are not significant for the concept maps. All in all, we may come to the hypothesis that the digital format motivates students to add more information in their dictionary entries and concept maps, which, in turn, might lead to a deeper engagement with the topic. Thus, we conclude that students should be motivated and carefully instructed to use digital tools for creating dictionary entries or concept maps if the goal is to motivate them for looking deeper into the domain.

Table 3: Results of the t-tests for digital and hand-written submissions. For each item, we give the arithmetic mean as well as the significance

	Digital	**Written**	**Significance**
Dictionary articles: points	16.0	10.5	n.s.
Dictionary articles: relations	5.97	2.90	$p<0.01$
Dictionary articles: items	22.6	10.2	$p<0.01$
Concept maps: points	17.56	15.24	n.s.
Concept maps: relations	3.88	4.12	n.s.
Concept maps: items	21.0	15.41	n.s.

This paper is part of the publication: Despot, K. Š., Ostroški Anić, A., & Brač, I. (Eds.). (2024). *Lexicography and Semantics. Proceedings of the XXI EURALEX International Congress.* Institute for the Croatian Language.

7. Discussion

Based on our quantitative evaluations we can provide preliminary indications for answering our research questions. However, further investigation with a bigger data set is necessary to evaluate if these tendencies are robust.

While we cannot give a final answer to (RQ1) for the concept map tasks, we can say for the dictionary task that there is a tendency towards a connection between higher mathematical competence and more engagement in solving the task. However, we cannot say anything about the direction of this relation based on our survey setting. Also, external motivation factors could influence both aspects.

Concerning (RQ2), both kinds of tasks seem to yield similar results concerning the engagement, though the focus is a bit different. If the didactical focus is on finding more different kinds of relations between terms, we recommend choosing dictionary tasks.

With our results for (RQ3), we recommend letting students create both, concept maps and dictionary entries, with digital tools to gain a deeper engagement. The reason for that might be that the tools enable changes in the structure and, along with it, rethinking about the task more easily.

In terms of future work, we plan to expand our evaluation to the whole student data to see if our hypotheses still hold. Additionally, it would be interesting to investigate whether the results are transferable to other domains of the exact sciences as well or mainly concern the didactics of mathematics.

References

Bowker, L. (2003). Specialized lexicography and specialized dictionaries. In P. G. J. van Sterkenburg (Ed.), *A practical guide to lexicography.* (pp. 154–164). John Benjamins.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. Aufl.). L. Erlbaum Associates.

Fuertes-Olivera, P. A., & Tarp, S. (2014). *Theory and Practice of Specialised Online Dictionaries. Lexicography versus Terminology*. De Gruyter.

Humbley, J. (2017). Specialised dictionaries. In P. A. Fuertes-Olivera (Ed.), *The Routledge Handbook of Lexicography*. Taylor & Francis Group. http://ebookcentral.proquest.com/lib/ubhildesheim-ebooks/detail.action?docID=5088188

Koplenig, A., & Müller-Spitzer, C. (2014). Questions of design. In C. Müller-Spitzer (Ed.), *Using online dictionaries* (pp. 189–204). De Gruyter.

Kruse, T., Heid, U., & Girnat, B. (2023). Using lexicography for learning mathematics. In M. Medved', M. Měchura, C. Tiberius, I. Kosem, J. Kallas, M. Jakubíček, & S. Krek (Eds.), *Electronic lexicography in the 21st century (eLex 2023): Invisible Lexicography. Proceedings of*

This paper is part of the publication: Despot, K. Š., Ostroški Anić, A., & Brač, I. (Eds.). (2024). *Lexicography and Semantics. Proceedings of the XXI EURALEX International Congress.* Institute for the Croatian Language.

the eLex 2023 conference (pp. 466–475). https://elex.link/elex2023/wp-content/uploads/97.pdf

Měchura, M. (2017). Introducing Lexonomy: An open-source dictionary writingand publishing system. In I. Kosem, C. Tiberius, M. Jakubíček, J. Kallas, S. Krek, & V. Baisa (Eds.), *Electronic lexicography in the 21st century. Proceedings of eLex 2017 conference* (pp. 662–679).

Contact information

Theresa Kruse
University of Hildesheim
theresa.kruse@fau.de

Ulrich Heid
University of Hildesheim
heidul@uni-hildesheim.de

Barbara Schmidt-Thieme
University of Hildesheim
bschmidt-thieme@imai.uni-hildesheim.de