Computers and Education Open 7 (2024) 100198

Contents lists available at ScienceDirect

Computers and Education Open

journal homepage: www.sciencedirect.com/journal/computers-and-education-open

Evaluating the content structure of intelligent tutor systems—A psychological network analysis

Markus W.H. Spitzer [*,a], Lisa Bardach [b], Younes Strittmatter [c], Jennifer Meyer [d], Korbinian Moeller [e]

[a] Martin-Luther University Halle-Wittenberg, Halle, Germany
[b] University of Tübingen,Tübingen, Germany
[c] Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, Providence, USA
[d] Leibniz Institute for Science and Mathematics Education (IPN), Kiel, Germany
[e] Loughborough University, Loughborough, United Kingdom

ARTICLE INFO

Keywords:
process data
intelligent tutoring system
mathematics
psychological network analysis

ABSTRACT

The adoption of intelligent tutoring systems (ITSs) worldwide has led to a considerable accumulation of process data as students interact with different learning topics within these systems. Typically, these learning topics are structured within ITSs (e.g., the fraction topic includes subtopics such as a fraction number line subtopic). However, there is a lack of methods that offer quick, data-driven insights into the content structure of ITSs, particularly through easily accessible visualizations. Here, we applied psychological network analysis to process data (230,241 students; 5,365,932 problem sets) from an ITS for learning mathematics to explore performance interdependencies between 40 different subtopics. We argue that the visualization of these content interdependencies allows for a quick empirical evaluation of the validity of the existing structuring of the respective learning content. These insights allow for deriving recommendations concerning potential changes in the ITS structure and are thus highly valuable for ITS developers. Our results are also relevant for researchers as the interdependencies illustrated through psychological network analysis can contribute towards a better understanding of the interplay between mathematical skills. Together, our results indicate that psychological network analysis represents a valuable data-driven method to evaluate and optimize ITSs.

1. Introduction

Intelligent tutoring systems (ITSs) provide adaptive learning opportunities for students and typically record process data as students engage with the learning content provided by the system. The widespread use of ITSs has led to enormous piles of such process data (e.g., [1–4]). For instance, within the ITS *Bettermarks* 230,241 students worked through 5, 365,932 problem sets from 9 different topics on fifth grade mathematics, which can be further broken down into 40 subtopics (between January 1, 2016, and August 31, 2023, in Germany). This example further illustrates that the datasets from ITSs are not only large but also complex, as data at different levels of granularity exist (e.g., students' performance on topics and constituting subtopics).

ITS often structure their content reflecting a traditional curriculum as they comprise different learning topics hierarchically along grade levels (e.g., fifth-grade mathematics topics in Germany include whole number divisions and multiplications, as well as fractions). Traditional curricula are designed by domain experts who possess deep knowledge and expertise in the field of mathematics education [2,5]. However, to our knowledge, the validity of how a broad range of different topics and subtopics are structured is rarely evaluated empirically by data-driven methods after being implemented.

In this study, we utilize psychological network analysis [6–8] to evaluate the content structure of ITSs, aiming to introduce a novel data-driven method for validating and advancing ITSs within the field of learning analytics. Our empirical investigation focuses on an ITS designed for mathematics learning. However, the implications of our approach are much broader and extend to ITSs covering diverse learning domains, such as language acquisition. Therefore, our study offers a foundational framework for future research concerning the validation of ITSs' content structures across various domains and educational contexts.

* Corresponding author at: Department of Cognition and Digital Learning, Martin-Luther University Halle-Wittenberg, Germany
E-mail address: markus.spitzer@psych.uni-halle.de (M.W.H. Spitzer).

https://doi.org/10.1016/j.caeo.2024.100198
Received 2 January 2024; Received in revised form 13 May 2024; Accepted 13 June 2024
Available online 18 June 2024

In the following sections, we first describe ITSs and how their accumulating process data can be leveraged to gain insights into students' performance on mathematical topics. We then introduce psychological network analysis in general and its application on process data to evaluate the content structure of ITSs and outline implications for research and practice (e.g., the design and further development of ITSs).

1.1. Intelligent tutoring systems

ITSs are digital learning environments with the primary aim to aid students' learning process through personalized learning experiences; for instance, by offering tailored support, feedback, and additional learning opportunities through suggested revisions in case students struggle with specific topics [9]. ITSs typically log the data generated during students' learning process. This process data typically include (but is not limited to) students' performance on topics and respective subtopics (e.g., accuracy), which topics students worked through as well as the time and date for each worked-through (sub)topic. The process data collected by ITSs is essential for the system itself because students' past performance may determine future (personalized) learning content and immediate feedback [10–12].

The analysis of process data may also help researchers and software developers to gain more detailed insights into students' learning processes and areas in which students may face difficulties (e.g., [1,2,13,14]). This information, in turn, may be used to refine the system; for instance, by restructuring the learning content of specific topics in a way that students face fewer difficulties during their learning process (e.g., [11,15]).

The use of process data also enables researchers to address important research questions on the relationships between different mathematical skills [16]. For instance, process data can be analyzed to test whether students' performance on a specific topic (e.g., arithmetic) predicts their performance on another later and more advanced topic (e.g., fractions). This allows for evaluating whether and how different topics within ITSs targeting specific skills are related, and, if analyzed longitudinally, how they build upon each other [16]. These insights may further be used to, for example, explicitly test whether students who revise problems from earlier topics known to predict later performance face fewer difficulties after such revisions [17].

As such, process data from ITSs serve as a valuable resource for (i) evaluating the general content structure of ITSs and (ii) advancing our understanding of mathematical skills more generally. Importantly, these insights may feed back into (iii) optimizing the respective ITSs [1,2]. However, to maximize the potential gain of evaluating process data, the application of methods that allow to quickly identify structures in exceptionally rich and complex process data are needed. This is where psychological network analysis comes into play.

1.2. Psychological networks analysis

Generally, psychological network analysis is a method to visualize correlation matrices. Variables are depicted as nodes and edges between nodes represent the correlations between variables [6,8,18], with the width of edges reflecting the strength of the correlation between two nodes. Force-distance algorithms can be applied to layout psychological networks in a way that distances between nodes, in addition to edge-widths, indicate how strongly nodes correlate with each other. For example, the Fruchterman-Reingold force-distance algorithm depicts nodes with relatively high correlations closer together, whereas nodes with relatively low correlations are pushed apart [19]. As a result, variables with relatively high intercorrelations are visualized as clusters, while variables with relatively low intercorrelations are pushed apart.

Importantly, other types of network analyses exist and have been applied to address research questions within the realm of learning analytics, such as social network analysis (e.g., [20–25]), semantic network analysis (e.g., [26]), or epistemic network analysis (e.g., [27]). The difference between these network analysis and psychological network analysis is that the edges within psychological network analysis are *estimated* (i.e., they represent the correlation between two variables) and are not directly obtained from the data (e.g., edges within social network analysis may represent the amount of communication between people and are thus directly obtained from the data).

Another important aspect of psychological network analysis is that large datasets are recommended as small samples, especially in combination with a complex data structure involving many variables, can negatively affect the accuracy of estimated edges (i.e., estimated correlations)[6]. Yet, as noted above, process data from ITSs typically comprises large sample sizes and thus, the application of psychological network analysis seems particularly well-suited for such datasets. However, despite the promise of psychological network analysis, their application to ITSs data is still scarce.

1.3. The application of psychological networks analysis to evaluate and explore process data from ITSs

In the following, we describe examples on how the application of psychological network analysis can serve as a powerful method to evaluate and explore complex interdependencies across different topics and subtopics implemented within ITSs. In our examples we specifically focus on an ITS developed for mathematics learning. However, psychological network analysis is not bound to any particular content. Thus, psychological network analysis can also be applied to other types of ITSs containing process data.

ITSs for mathematics learning typically comprise several different topics, each of which contains multiple subtopics. For instance, within the ITS *Bettermarks* (see Methods), there is a *Basics of Fractions* topic which includes the subtopics *1. Shares of a whole, 2. Expand, shorten, and compare fractions, 3. Fractions as quotients, 4. Fractions on the number line, 5. Shares of sizes*, and *6. Fractions and percentages*. Students' performance on these subtopics is presumed to reflect students' fraction understanding. Accordingly, one would expect that students who perform well on one fraction subtopic should also perform well on another fraction subtopic and vice versa. In other words, one would expect high intercorrelations between fraction subtopics, as they were designed to measure the same underlying concept—fraction understanding. When applying psychological network analysis, relatively high intercorrelations between subtopics within each topic would be reflected as clusters of subtopics arranged relatively close to each other. In case such clusters could be observed for each topic, this would substantiate the validity of the respective content structure (i.e., subtopics meant to constitute a topic are actually related [28]).

Moreover, a large body of research suggests that students' mathematics achievements are relatively robust across their educational path [29–36]. In other words, a student who performs well on one mathematical topic is likely to also perform well on another mathematical topic. In contrast, a student who faces difficulties with one topic is also likely to face difficulties with another topic. For instance, evidence from several studies suggested that students whole number arithmetic skills and whole number knowledge predict their fraction understanding, indicating that proficient whole number processing serves as a building block for later more complex mathematical concepts such as fraction understanding [31–33]. Similar results may also be observed for other formats of rational numbers, such as fractions, decimals, and percentages [37]. However, empirical evidence substantiating relatively high interdependencies between the three formats of rational numbers is missing (for associations of fractions and decimals see [38–40]).

Nevertheless, some topics may build upon each other more strongly than others, and variations in interdependencies between topics should be observed when exploring data on students' performance across a range of different topics. In addition, one may also observe that some topics correlate relatively highly with many other topics, whereas some topics rather correlate highly with just a few other topics, if any. The application of psychological network analysis may be promising for such a use case as it allows to quickly visualize correlations between topics in a user-friendly and intuitive way to better understand how different topics are related to each other.

In sum, when applying psychological network analysis to process data of ITSs for learning mathematics, one may expect the following observations. First, subtopics of the same topic that are supposed to measure the same underlying construct (e.g., fraction understanding), should correlate highly and thus should be visualized as clusters within the specified psychological networks. Moreover, when a set of different constructs (i.e., topics) was implemented in the ITS, one may expect distinct clusters for each construct/topic. Finally, one may also expect that two related constructs that build upon each other (e.g., different rational number formats) should be positioned relatively close to each other within the resulting psychological network.

1.4. Study overview

In this study, we applied psychological network analysis to explore interdependencies between students' performance on 40 different subtopics stemming from 9 different fifth-grade topics. Therefore, we leveraged data from the ITS *Bettermarks*. We conducted an exemplary psychological network analysis as a use case on how the application of this method may help to identify the structure of students' performances across and within topics. We expected to observe subtopics constituting a topic to cluster together. We also expected topics which target similar or related competencies to be closely positioned, whereas unrelated topics should be further apart from each other (for a visualization of our procedure and potential implications, see Fig. 1).

2. Methods

2.1. The ITS bettermarks

Bettermarks is an ITS with a primary focus on mathematics learning (e.g., [3,41–44]), which was designed to cater to students across diverse age groups and educational levels (age-range 9-18, grades 4-12). Within *Bettermarks*, teachers can assign problem sets from different topics to their students to work through in class or at home. The ITS comprises 31 different mathematical topics suited for fifth graders (see Fig. 2). Each of these topics is further subdivided into subtopics which further consist of several problem sets. Each problem set comprises an average of nine problems. The ITS logs the IDs of topics, subtopics, and problem sets as well as the average accuracy of students on these problem sets together with the date and time when students worked-through problem sets.

Bettermarks incorporates several adaptive features to enhance the learning experience. First, students and teachers receive immediate feedback on an average performance score for each assigned problem set. This feature enables continuous monitoring of progress for both students and teachers. Second, the platform offers motivational incentives, allowing students to earn stars and coins based on their performance. Stars are awarded for achieving 100% accuracy, while coins are earned within specific accuracy ranges (e.g., 90-99% = 3 coins; 75-89% = 2 coins; 60-74% = 1 coin). Third, students can repeat problem sets as often as they wish. However, the parameters of the problem sets change with each attempt, discouraging rote memorization and promoting active problem-solving. Fourth, if students face difficulties with a topic, the system automatically recommends revision problem sets to students to close students' knowledge gaps.

As noted above, problem sets comprise several problems which may require one or more steps to solve. After each step, students receive immediate feedback on whether their solution was correct or not. This feedback is tailored to the specific content and errors made by the student, providing guidance for improvement. In case students make a mistake on one of the steps, the correct answer is not provided, but students get a second chance to complete this step and to learn from their mistakes with the help of feedback. The accuracy of problem sets decreases with each mistake students make and thus students are motivated to use only one attempt on each problem.

Bettermarks shares its anonymous data with researchers for secondary data analysis on request. Sensitive personal information (e.g., age or

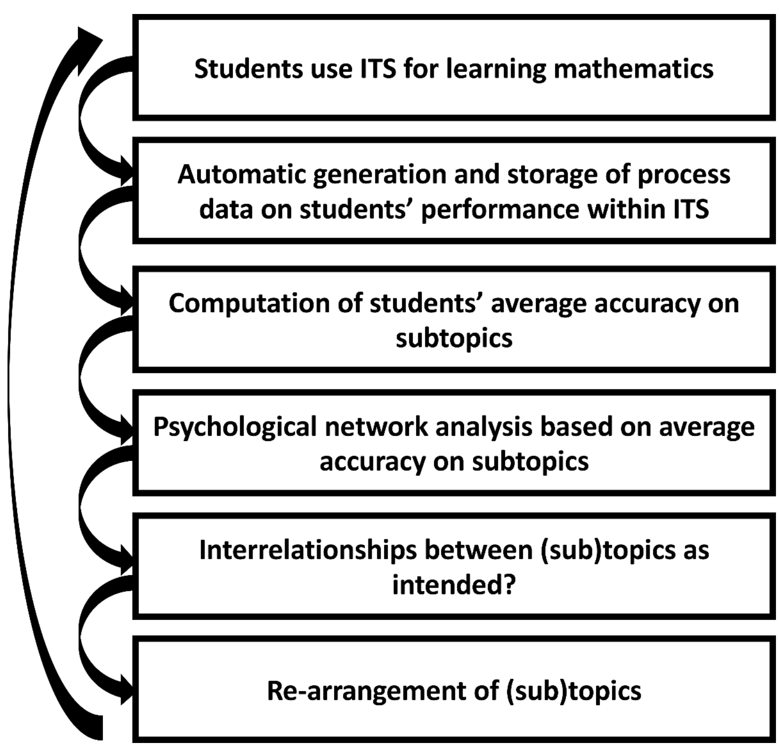

Fig. 1. Process diagram displaying the steps of the study and potential implications.

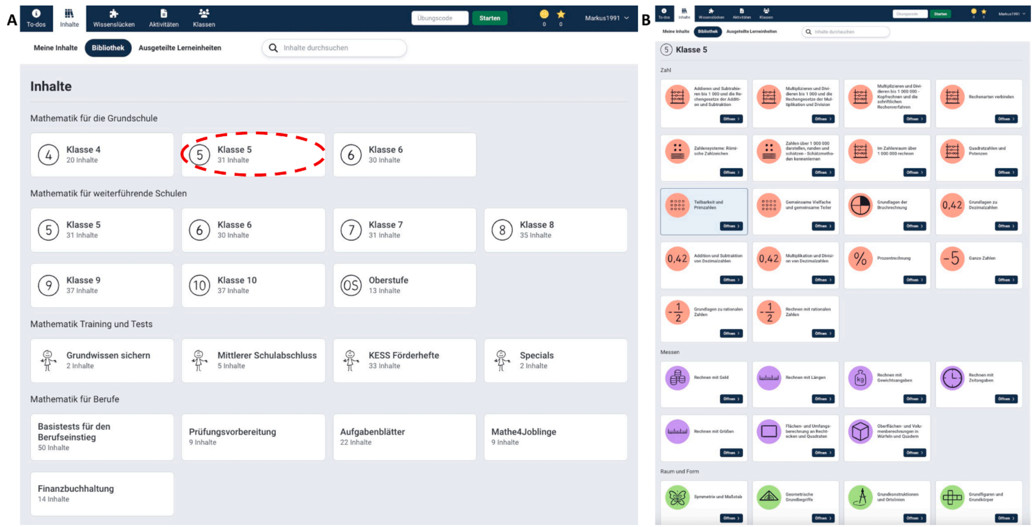

Fig. 2. The user interface of the *Bettermarks* system used in Germany. A: Grade levels can be selected and for each grade level several different topics exist. B: The selection of the topics suited for fifth-grade students which appears after selecting grade 5 (Klasse 5). Red dashed ellipse highlights the grade level.

gender) about users is not available as the data is fully anonymous. It is crucial to note that *Bettermarks* played no role in the study's design or outcomes, maintaining complete independence from the investigation. Thus, the study's findings may not necessarily reflect the views or opinions of *Bettermarks*.

2.2. *The present dataset*

The dataset for this study was obtained from *Bettermarks* based on the following criteria. First, we considered the topics suited for grade five implemented in the German *Bettermarks* system. Second, we included all problem sets from these topics that were computed between August 1st, 2016, and August 30th, 2023. Third, we considered students best results for any given problem set they computed as a proxy for their performance on this problem set. However, repetition rates for problem sets were low (average 1.4) and results did not differ substantially when only considering students first attempts. Fourth, we only included problem sets that were computed by a considerable number of students (>1000 students computed each considered problem set) to obtain robust accuracy estimates on students' performance. Fifth, each student computed at least 10 problem sets to be included. Finally, we only considered topics which were worked through by over 40,000 students to obtain robust accuracy estimates on topics. Based on these criteria, our dataset comprised 230,241 students who worked through 5,365,932 problem sets from 9 topics separated into 40 subtopics.

2.3. *Data analysis*

Our data analysis was conducted with the R software [45]. We applied the igraph package for our psychological network analysis with marginal correlations [46]. All marginal correlations were Pearson correlations.

We applied marginal correlations and not partial correlations because not all students worked through all topics and subtopics. As this leads to missing values for topics and subtopics, the use of marginal correlations between subtopics allowed us to base our evaluation on as much data as possible. However, marginal correlations reflect the association between two variables without considering the influence of other variables (see also the discussion of limitations in our discussion section).

The psychological network analysis drew on a correlation matrix which was based on students' average performance on the considered subtopics. The psychological network analysis was conducted using the igraph package in R that allows to create the psychological network analysis plots within R based on the correlation matrix (see https://osf.io/mn469/ for the data analysis scripts as well as the data for carrying out the psychological network analysis). The average performance of each student on each subtopic was reflected by the average accuracy on all worked-through problem sets per subtopic. For instance, if a student worked through three problem sets of a subtopic and achieved accuracies of 70%, 80%, and 90%, then the average subtopic accuracy of 80% resulted from the average of three problem set accuracies.

3. Results

The results from the psychological network analysis are depicted in Fig. 3. Each node depicts a subtopic with a topic abbreviation and a number that indicates the sequential appearance of the subtopics within *Bettermarks*. Table 1 lists all 9 topics including the topic abbreviation, the topic name, the average accuracy of each topic, and the number of students who worked-through each topic (n).

Our psychological network analysis provides an overview on the interdependencies between students' performance on 9 different topics (see Fig. 3). In the following two sections, we describe how this visualization can be used to address different kinds of research questions.

3.1. *Evaluating the validity of the content structure*

The psychological network analysis indicated that the subtopics of most topics were located relatively closely together, indicating high intercorrelations between subtopics of the same topic. For instance, we identified relatively tight clusters for all subtopics of the (i) *Percentages* topic (Per; purple nodes in Fig. 3), (ii) *Multiplication and Division of Decimal Numbers* topic (Dec-Mul; yellow nodes), (iii) *Basic of Decimals* topic (Dec; orange nodes), (iv) *Calculating with Lengths* topic (Len; dark-blue nodes), (v) *Basic Figures and Basic Solids* topic (Fig; green nodes), and (vi) *Area and Perimeter Calculation on Rectangles and Squares* topic (Are; red nodes). However, the cluster of *Area and Perimeter Calculation on Rectangles and Squares* subtopics (Are; red nodes) comprised one subtopic of the *Basic Geometrical Concepts* topic (Geo-3; light-blue nodes). The subtopics of this *Basic Geometrical Concepts* topic (Geo; light-blue nodes), in contrast, were rather diffused, though still connected with edges. Subtopics of the *Whole Numbers* topic (Who; magenta nodes) were visualized in a straight line between *Percentages* subtopics and subtopics from the two topics on decimals (Dec and Dec-Mul). Finally, all subtopics of the *Basics of Fractions* were relatively close

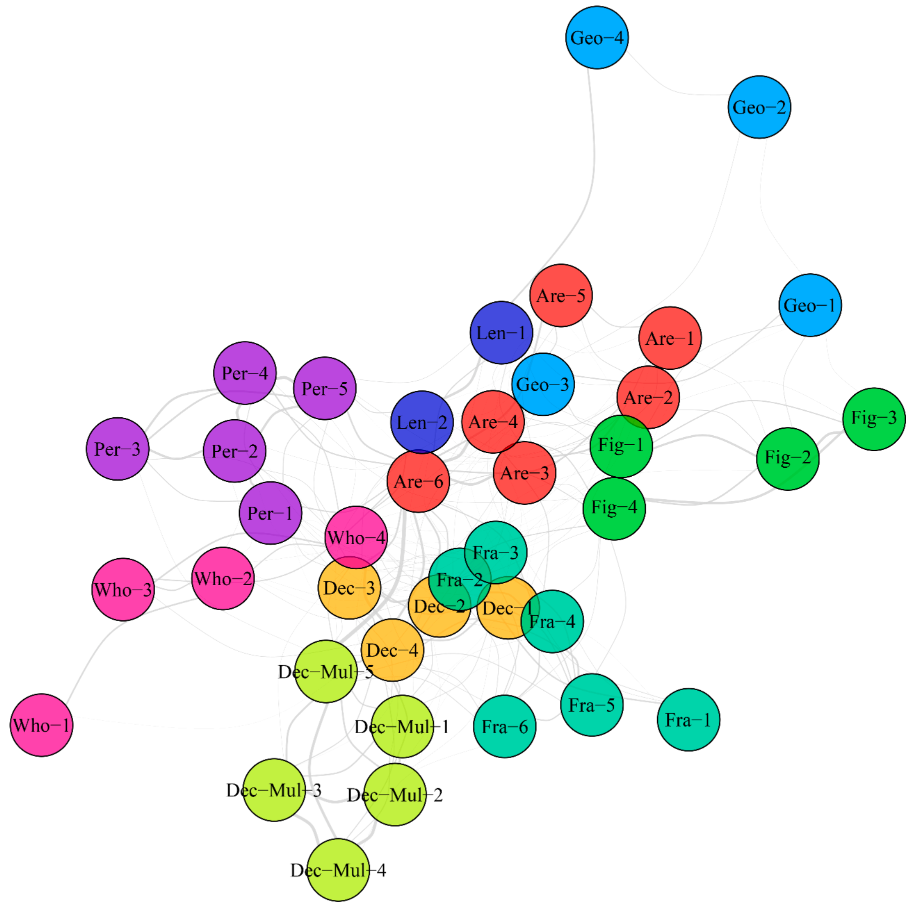

Fig. 3. Psychological network depicting interdependencies between grade-five topics and subtopics. This network depicts all edges with a minimum partial correlation of r =.3. Edge width scales with correlation strength. Node colors represent topics.

Table 1
Topic ID, Topic Name, Average Accuracy, Standard Deviation (SD), and n Problem sets.

ID	Topic	Accuracy	SD	n
Are	Area and Perimeter Calculation on Rectangles and Squares	.81	.27	638,479
Dec	Basic of Decimals	.86	.23	471,256
Dec-Mul	Multiplication and Division of Decimal Numbers	.84	.26	406,171
Fig	Basic Figures and Basic Solids	.78	.28	383,252
Fra	Basics of Fractions	.88	.22	1,180,959
Geo	Basic Geometrical Concepts	.86	.23	647,802
Len	Calculating with Lengths	.80	.26	248,199
Per	Percentages	.84	.24	830,144
Who	Whole Numbers	.86	.22	559,670

together, but the three subtopics Fra-2, Fra-3, and Fra-4 were also very closely tied to the two subtopics Dec-1 and Dec-2 of the *Basics of Decimals topic*. Taken together, these insights help to make quick judgements on whether subtopics of the same topic form the expected clusters, thus providing evidence for the validity of the content structure of the ITS.

3.2. Psychological network analysis as an exploration lens for addressing research questions on content-specific learning

When evaluating the position of subtopics across topics, our psychological network analysis showed that topics known to build upon each other over the course of students' development of mathematics competences were in fact estimated to be relatively closely together. For instance, the observation of relatively high interdependencies between three specific fraction subtopics and two specific decimal subtopics could hint towards the idea that these two topics may build on similar underlying constructs (e.g., understanding rational numbers). This is further substantiated by the observation that subtopics of both decimal topics and the fraction topic were relatively closely positioned. Thus, these results indicate that the illustration of interdependencies via psychological networks can contribute towards a better understanding of the interrelations between mathematical skills.

4. Discussion

In this article, we applied psychological network analysis to empirically evaluate and visualize the content structure of an ITS for learning mathematics reflected by interdependencies of students' performance on different (sub)topics (e.g., a fraction number line subtopic within a fraction topic). We demonstrated how psychological networks can be used to better understand complex interdependencies between many different variables. This information is useful for addressing different questions—from validating the content structure of ITSs to the interplay between different mathematical skills.

Our example considered fifth-grade mathematics topics and their subtopics taken from *Bettermarks* (based on data from 230,241 students students working on 5,365,932 problem sets) and showed how psychological network analysis can provide an easily accessible overview of performance interdependencies between the different topics and subtopics. With respect to the evaluation of the content structure of ITSs, our psychological network analysis indicated that subtopics of the same topic indeed clustered together, indicating relatively higher intercorrelations between these subtopics than with subtopics of other topics. Thereby, psychological network analysis provided quick empirical evidence by means of user-friendly visualizations, revealing that subtopics that are meant to measure similar constructs (e.g., fraction understanding) actually seem to measure something similar. Our example also suggests that psychological network analysis allows for gaining deeper insights into the specific structure of interdependencies between different topics. For instance, we observed that (sub)topics on fractions, decimals, and percentages were arranged relatively close together, potentially reflecting the underlying ability of understanding rational numbers.

Taken together, these results highlight the promise of applying psychological network analysis as a method for evaluating the validity of the structuring of learning content in ITSs. As the method is generic and data driven, it should be easily transferable and applicable to other large-scale data sets from ITSs. With more and more datasets from different ITSs now being publicly available with free access for researchers (e.g., [1,47]), facing such large datasets for the first time can be overwhelming. As observed in the present study, psychological network analysis represents a valuable instrument to visualize interdependencies between a range of variables to help to identify and understand the underlying data structure of such ITSs.

Psychological network analysis is not only interesting for researchers, but also for software developers working on optimizing ITSs to foster the learning progress of students. Most ITSs incorporate personalized adaptive recommendations for students to revise (sub)topics to close identified knowledge gaps based on the idea that students will face fewer difficulties with the current (sub)topic after such revisions. In this context, the application of psychological network analysis provides a data-driven and interpretable method that helps to gain insights into specific interdependencies between students performance on specific topics and their associated subtopics. This information is highly relevant for optimizing adaptive recommendations for students to revise specific (sub)topics to close their knowledge gaps. The evaluation of process data with psychological network analysis may also help software developers to optimize the structuring of (sub)topics in ITSs. For instance, software developers may consider to restructure certain (sub)topics based on the gained information on which subtopics do not fit in expected cluster. The restructured (sub)topic could then be re-evaluated using psychological network analyses, providing insights into how the clustering might change with the new ordering of (sub)topics.

4.1. Limitations and future directions

The current results should be interpreted in light of several limitations, yet these limitations also point towards promising future research directions. One notable limitation is our utilization of marginal correlations in conducting psychological network analysis. This approach did not account for potential influences of other (sub)topics on the clustering of topics and their interrelations. The choice of marginal correlations was motivated by the aim to incorporate as much data as possible, including cases with missing data. Nonetheless, future studies might explore the application of psychological network analysis for process data from ITSs with partial correlations, allowing to control for the influence of other variables [6]. Applying partial correlations also allows to consider recent bootstrapping techniques so that a 95% confidence interval can be estimated for all edges, which further increases scientific rigor [7].

Another limitation is that this study evaluated the correlations between mathematical (sub)topics without controlling for differences in the times when specific problem sets were worked through. In light of this limitation, we propose that psychological network analysis may well be applied as a first step to evaluate interdependencies (correlations) between a wide range of (sub)topics, but then needs to be followed up by additional analyses to understand developmental trajectories of learning (mathematics) in more detail. For instance, a promising avenue for future research could involve examining the direction of associations between (sub)topics by explicitly considering the temporal sequence in which students worked through them. In this study, our primary objective was to demonstrate the value of psychological network analysis as an exploration lens, which provides a rapid overview of interdependencies between a rich set of variables as a first instance of data analysis.

Moreover, we applied psychological network analysis to fifth-grade mathematical problems as an example. Future work that adopts psychological network analysis to investigate interdependencies across mathematical subtopics of higher or lower grades is needed to evaluate its suitability for identifying interdependencies of mathematics (sub)topics in other samples and age groups. Furthermore, future studies should apply psychological network analysis to investigate interdependencies for other subject domains than mathematics as well as across subject domains (e.g., How is students' performance in mathematics, physics, native and foreign language classes, music, etc. associated?).

Another interesting avenue for future research is to determine whether changes to the content structure based on a psychological network analysis lead to changes in the identified structure (i.e., a more consistent clustering of subtopics within the psychological network). In other words, one may investigate whether a modification to the content structure of subtopics, as suggested by a psychological network analysis, leads to the expected outcome of subtopics of the same topics clustering closer together relative to other subtopics of other topics in a subsequent data collection. This may also include the consideration of other variables than students' performance, such as time required for working through problems.

4.2. Conclusion

In conclusion, the present example clearly indicates that psychological network analysis can be a powerful method to quickly visualize, evaluate, and explore interdependencies between a range of different variables (e.g., subtopics) within an ITS. This information can be used to evaluate the validity of the existing structure of the learning content in the respective ITS and may also provide valuable insights into interrelations between mathematical skills. At the same time, the insights from psychological network analysis can lead to improvements of ITSs. For instance, based on the observed interdependencies between subtopics, recommendations regarding which (sub)topics to revise when difficulties are encountered and recommendations regarding adaptations to the structuring of the learning content in general can be derived. As such, psychological network analysis seems to be a useful method for (educational) researchers and software developers alike, who are both interested in better understanding how to best aid students' learning.

CRediT authorship contribution statement

Markus W.H. Spitzer: Writing – original draft, Visualization, Data curation, Conceptualization. **Lisa Bardach:** Writing – review & editing, Writing – original draft, Conceptualization. **Younes Strittmatter:** Writing – review & editing, Conceptualization. **Jennifer Meyer:** Writing – review & editing, Conceptualization. **Korbinian Moeller:** Writing – review & editing, Conceptualization.

Declaration of competing interest

There are no known conflicts of interest associated with this publication.

References

[1] Koedinger KR, Carvalho PF, Liu R, Mclaughlin EA. An astonishing regularity in student learning rate. Proc Natl Acad Sci 2023;120(13). https://doi.org/10.1073/pnas.e2221311120

[2] Ritter S, Anderson JR, Koedinger KR, Corbett A. Cognitive tutor : applied research in mathematics education. Psychonomic Bull Rev 2007;14(2):249–55.

[3] Spitzer MWH, Moeller K, Musslick S. Assignment strategies modulate students academic performance in an online learning environment during the first and second covid-19 related school closures. Plos One 2023;18(5):e0284868.

[4] Meeter M. Primary school mathematics during Covid-19: no evidence of learning gaps in adaptive practicing results. Trend Neurosci Educ 2021;25(October): 100163. https://doi.org/10.1016/j.tine.2021.100163.

[5] Anderson JR, Corbett AT, Koedinger KR, Pelletier R. Cognitive tutors: lessons learned. J Learn Sci 1995;4(2):167–207.

[6] Epskamp S, Borsboom D, Fried EI. Estimating psychological networks and their accuracy: a tutorial paper. Behav Res Method 2018;50(1):195–212. https://doi.org/10.3758/s13428-017-0862-1.

[7] Epskamp S, Fried EI. A tutorial on regularized partial correlation networks. Psychol Methods 2018;23(4):617–34. https://doi.org/10.1037/met0000167.

[8] Borsboom D, Cramer AO. Network analysis: an integrative approach to the structure of psychopathology. Annual Rev Clin Psychol 2013;9:91–121. https://doi.org/10.1146/annurev-clinpsy-050212-185608.

[9] VanLehn K. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. Educ Psychol 2011;46(4):197–221. https://doi.org/10.1080/00461520.2011.611369.

[10] Mavrikis M, Holmes W. Intelligent learning environments: design, usage and analytics for future schools. Shap Future Schools Digit Technol: Int Handbook 2019:57–73.

[11] Mavrikis M, Rummel N, Wiedmann M, Loibl K, Holmes W. Combining exploratory learning with structured practice educational technologies to foster both conceptual and procedural fractions knowledge. Educ Technol Res Dev 2022;70(3): 691–712.

[12] Mousavinasab E, Zarifsanaiey N, R. Niakan Kalhori S, Rakhshan M, Keikha L, Ghazi Saeedi M. Intelligent tutoring systems: a systematic review of characteristics, applications, and evaluation methods. Interact Learn Environ 2021;29(1):142–63.

[13] Rittle-Johnson B, Koedinger K. Iterating between lessons on concepts and procedures can improve mathematics knowledge. Brit J Educ Psychol 2009;79(3): 483–500. https://doi.org/10.1348/000709908X398106.

[14] Rau MA, Aleven V, Rummel N. Intelligent tutoring systems with multiple representations and self-explanation prompts support learning of fractions. Front Artif Intell Appl 2009;200(1):441–8. https://doi.org/10.3233/978-1-60750-028-5-441.

[15] Rittle-Johnson B, Koedinger KR. Designing knowledge scaffolds to support mathematical problem solving. Cognit Instruct 2005;23(3):313–49. https://doi.org/10.1207/s1532690xci2303_1.

[16] Spitzer MWH, Moeller K. Predicting fraction and algebra achievements online: a large-scale longitudinal study using data from an online learning environment. J Comput Assist Learn 2022. https://doi.org/10.31234/osf.io/rw6b9.

[17] Wortha SM, Klein E, Lambert K, Dackermann T, Moeller K. The relevance of basic numerical skills for fraction processing: Evidence from cross-sectional data. Plos One 2023;18(1):e0281241.

[18] Cramer AO, Waldorp LJ, Van Der Maas HL, Borsboom D. Comorbidity: a network perspective. Behav Brain Sci 2010;33(2-3):137–50. https://doi.org/10.1017/S0140525X09991567.

[19] Fruchterman TMJ, Reingold EM. Graph drawing by force-directed placement. Softw-Practice Exper 1991;2(1):387–1129–1164. https://doi.org/10.1007/978-3-319-64471-4_31.

[20] Norz LM, Dornauer V, Hackl WO, Ammenwerth E. Measuring social presence in online-based learning: An exploratory path analysis using log data and social network analysis. Internet Higher Educ 2023;56(October 2022):100894. https://doi.org/10.1016/j.iheduc.2022.100894.

[21] Dawson S, Gašević D, Siemens G, Joksimovic S. Current state and future trends: a citation network analysis of the learning analytics field. ACM Int Conf Proc Ser 2014:231–40. https://doi.org/10.1145/2567574.2567585.

[22] Saqr M, López-Pernas S. How CSCL roles emerge, persist, transition, and evolve over time: a four-year longitudinal study. Comput Educ 2022;189(June). https://doi.org/10.1016/j.compedu.2022.104581.

[23] Stasewitsch E, Barthauer L, Kauffeld S. Knowledge transfer in a two-mode network between higher education teachers and their innovative teaching projects. J Learn Anal 2022;9(1):93–110. https://doi.org/10.18608/jla.2022.7427.

[24] Mallavarapu A, Lyons L, Uzzo S. Exploring the utility of social-network-derived collaborative opportunity temperature readings for informing design and research of large-group immersive learning environments. J Learn Anal 2022;9(1):53–76. https://doi.org/10.18608/jla.2022.7419.

[25] Chen B, Poquet O. Networks in learning analytics: where theory, methodology, and practice intersect. J Learn Anal 2022;9(1):1–12. https://doi.org/10.18608/jla.2022.7697.

[26] Chowdhury M, Dixon LQ, Kuo L-J, Donaldson JP, Eslami Z, Viruru R, Luo W. Digital game-based language learning for vocabulary development. Comput Educ Open 2024;6:100160.

[27] Shaffer DW, Collier W, Ruis AR. A tutorial on epistemic network analysis: analyzing the structure of connections in cognitive, social, and interaction data. J Learn Anal 2016;3(3):9–45. https://doi.org/10.18608/jla.2016.33.3.

[28] Borsboom D, Mellenbergh GJ, Van Heerden J. The concept of validity. Psychol Rev 2004;111(4):1061–71. https://doi.org/10.1037/0033-295X.111.4.1061.

[29] Siegler RS, Duncan GJ, Davis-Kean PE, Duckworth K, Claessens A, Engel M, Susperreguy MI, Chen M. Early Predictors of High School Mathematics Achievement. Psychol Sci 2012;23(7):691–7. https://doi.org/10.1177/0956797612440101.

[30] Murnane RJ, Willett JB, Levy F. The growing importance of cognitive skills in wage determination. Rev Econ Stat 1995;77(2):251–66.

[31] Jordan NC, Hansen N, Fuchs LS, Siegler RS, Gersten R, Micklos D. Developmental predictors of fraction concepts and procedures. J Exp Child Psychol 2013;116(1): 45–58. https://doi.org/10.1016/j.jecp.2013.02.001.

[32] Hansen N, Jordan NC, Fernandez E, Siegler RS, Fuchs L, Gersten R, Micklos D. General and math-specific predictors of sixth-graders' knowledge of fractions. Cognit Dev 2015;35:34–49. https://doi.org/10.1016/j.cogdev.2015.02.001.

[33] Vukovic RK, Fuchs LS, Geary DC, Jordan NC, Gersten R, Siegler RS. Sources of individual differences in children's understanding of fractions. Child Dev 2014;85 (4):1461–76. https://doi.org/10.1111/cdev.12218.

[34] National Mathematics Advisory Panel. Foundations for success: the final report of the National Mathematics Advisory Panel. Tech. Rep.. 2008.

[35] Bailey DH, Hoard MK, Nugent L, Geary DC. Competence with fractions predicts gains in mathematics achievement. J Exp Child Psychol 2012;113(3):447–55. https://doi.org/10.1016/j.jecp.2012.06.004.

[36] Bailey DH, Siegler RS, Geary DC. Early predictors of middle school fraction knowledge. Dev Sci 2014;17(5):775–85. https://doi.org/10.1111/desc.12155.

[37] Tian J, Siegler RS. Which type of rational numbers should students learn first? Educ Psychol Rev 2018;30(2):351–72. https://doi.org/10.1007/s10648-017-9417-3.

[38] Braithwaite DW, Siegler RS. A unified model of arithmetic with whole numbers, fractions, and decimals. Psychol Rev 2023;(July). https://doi.org/10.1037/rev0000440.

[39] Braithwaite DW, McMullen J, Hurst MA. Cross-notation knowledge of fractions and decimals. J Exp Child Psychol 2022;213:105210. https://doi.org/10.1016/j.jecp.2021.105210.

[40] Braithwaite DW, Sprague L, Siegler RS. Toward a unified theory of rational number arithmetic. J Exp Psychol: Learn Memory Cognit 2021;48(10):1470–83. https://doi.org/10.1037/xlm0001073.

[41] Stapel M, Zheng Z, Pinkwart N. An ensemble method to predict student performance in an online math learning environment. Proceedings of the 9th International Conference on Educational Data Mining. 2016. p. 231–8.

[42] Whalen KA, Renkl A, Eitel A, Glogger-Frey I. Digital re-attributional feedback in high school mathematics education and its effect on motivation and achievement. J Comput Assist Learn 2024;40(2):478–93.

[43] Scharnagl S, Evanschitzky P, Streb J, Spitzer M, Hille K. Sixth graders benefit from educational software when learning about fractions: a controlled classroom study. Numeracy 2014;7(1):4.

[44] Spitzer MWH, Moeller K. Performance increases in mathematics within an intelligent tutoring system during covid-19 related school closures: a large-scale longitudinal evaluation. Comput Educ Open 2024;6:100162.

[45] R Core Team. R: A Language and Environment for Statistical Computing. 2013. http://www.r-project.org/.

[46] Csardi G, Tamas N. The igraph software package for complex network research. InterJournal 2006;Complex Sy:2006.

[47] Koedinger KR, Baker RS, Cunningham K, Skogsholm A, Leber B, Stamper J. A data repository for the EDM community: the PSLC datashop. 2010. https://doi.org/10.1201/b10274.ISBN 9781439804582