The potential contributions of concept maps for learning website to assessment for learning practices

Published in:
Knowledge Management and E-Learning

Document Version:
Publisher's PDF, also known as Version of record

Queen's University Belfast - Research Portal:
[Link to publication record in Queen's University Belfast Research Portal](#)

Knowledge Management & E-Learning
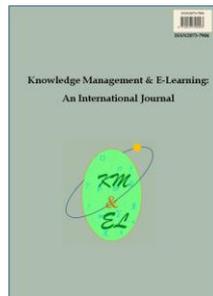
The potential contributions of concept maps for learning website to assessment for learning practices

Mehmet Filiz
David L. Trumpower
Shehzad Ghani
Sait Atas
Arun Vanapalli
The University of Ottawa, Ottawa, ON, Canada

The potential contributions of concept maps for learning website to assessment for learning practices

Mehmet Filiz*

Faculty of Education
The University of Ottawa, Ottawa, ON, Canada
E-mail: mehmetfiliz52@hotmail.com

David L. Trumpower

Faculty of Education
The University of Ottawa, Ottawa, ON, Canada
E-mail: david.trumpower@uottawa.ca

Shehzad Ghani

Faculty of Education
The University of Ottawa, Ottawa, ON, Canada
E-mail: sghani@uottawa.ca

Sait Atas

Faculty of Education
The University of Ottawa, Ottawa, ON, Canada
E-mail: sait.atas@uottawa.ca

Arun Vanapalli

Faculty of Education
The University of Ottawa, Ottawa, ON, Canada
E-mail: vanapalliarun@gmail.com

*Corresponding author

Abstract: The purpose of this paper is to examine the promising contributions of the Concept Maps for Learning (CMfL) website to assessment for learning practices. The CMfL website generates concept maps from relatedness degree of concepts pairs through the Pathfinder Scaling Algorithm. This website also confirms the established principles of effective assessment for learning, for it is capable of automatically assessing students' higher order knowledge, simultaneously identifying strengths and weaknesses, immediately providing useful feedback and being user-friendly. According to the default assessment plan, students first create concept maps on a particular subject and then they are given individualized visual feedback followed by associated instructional material (e.g., videos, website links, examples, problems, etc.) based on a comparison of their concept map and a subject matter expert's map. After studying the feedback and instructional material, teachers can monitor their

students' progress by having them create revised concept maps. Therefore, we claim that the CMfL website may reduce the workload of teachers as well as provide immediate and delayed feedback on the weaknesses of students in different forms such as graphical and multimedia. For the following study, we will examine whether these promising contributions to assessment for learning are valid in a variety of subjects.

Biographical notes: Mehmet Filiz is currently a third year Ph.D. student of Teaching, Learning and Evaluation in the Faculty of Education at the University of Ottawa, working under the supervision of Professor Trumpower. His doctoral work focuses on the use-ability and effectiveness of an online concept mapping tool for formative assessment. He holds an M.A. in Mathematics Education from Karadeniz Technical University, Turkey. His Ph.D. studies are sponsored by The Turkish Ministry of National Education.

David L. Trumpower is an associate professor of Teaching, Learning, and Evaluation in the Faculty of Education at the University of Ottawa. One of his main research interests includes a focus on structural assessment methods, particularly in the areas of mathematics and the sciences. He is currently exploring online and formative applications of knowledge map assessment. Another of his main areas of research interest is statistics education. Professor Trumpower teaches courses on quantitative research methods and analysis, learning processes, foundations of assessment and assessment for learning.

Shehzad Ghani is a PhD candidate at the University of Ottawa's Faculty of Education. With a background in computer information systems and administration, his current interests are e-Learning, Educational Innovation, ICT Information & Communication Technologies, Learning and teaching in the online environment, and formative assessment. He is conducting research on learning management systems role in formative assessment.

Sait Atas is a PhD candidate at the University of Ottawa's Faculty of Education. His academic training is on Measurement, Evaluation, and Assessment in education. His main interests are statistics education, formative assessment, and concepts maps as a means of increasing deeper understanding in statistics. As a research team, we are using concepts maps as an assessment tool to visually represent students' and teachers' domain knowledge and then, using the results to identify misconceptions in statistics.

Arun Vanapalli has undergraduate degrees from the University of Ottawa in Engineering and Education. He is currently pursuing a Master's degree in Education while working for the Government of Canada. His goal is to help write and edit textbooks for high school level mathematics and science classes.

1. Introduction

Assessment is one of the crucial components of education (Gikandi, Marrow, & Davis, 2011), and is required for three broader aims (Black, 1993):

- The certification of individual student achievement;

- The accountability of educational institutions via the comparison of results;
- Direct assistance to learning through useful feedback.

1.1. Assessment for learning

Recently, there has been a shift from a primary focus on summative assessment to formative assessment (Irons, 2007; Greenstein, 2010), yet what formative assessment is remains an issue. Some claim that formative assessment refers to assessment instruments whereas others believe that it is a process (Bennett, 2011). Defining formative assessment as only an instrument causes confusion when formative feedback is given on summative assessment (Irons, 2007). For instance, $3^2 + 2 \times 4$ is an item which assesses students' understanding of the order of operations in mathematics. If the item is asked at the end of a unit, it has summative value. Provided that a student answered 44 rather than 17, it may be concluded that the student ignores the hierarchy of multiplication over addition. This information might have formative value if the teacher makes some instructional decisions to address this misconception (Good, 2011). Therefore, even though an assignment might be designed or planned as summative, its methodology, data analysis and use of results determine whether it is formative or summative (Dunn & Mulvenon, 2009). Consequently, the mistake of incorrectly concluding that teachers are able to infer students' specific misconceptions from every item of summative assessments might be made. Due to this issue, the supporters of the process view started using the term "Assessment for Learning" or "Formative Evaluation" instead of "Formative Assessment" in the literature (Bennett, 2011). With this regard, Cizek (2010) defines assessment for learning as:

> *"the collaborative processes engaged in by educators and students for the purpose of understanding the students' learning and conceptual organization, identification of strengths, diagnosis of weaknesses, areas for improvement, and as a source of information that teachers can use in instructional planning and students can use in deepening their understandings and improving their achievement (pp. 6-7)."*

1.2. The effectiveness of assessment for learning

Despite the challenges to assessment for learning including resources and time (Cizek, 2010), several studies reveal the positive impact of assessment for learning on students' achievement. For example, Black and Wiliam (1998) concluded that assessment for learning increases students' performance by reviewing around 250 articles based on assessment for learning. These studies are related to feedback, self-assessment, and peer assessment. Likewise, Nyquist (2003) came to a similar conclusion after conducting a meta-analysis on use of feedback for formative assessment purposes. And, in an empirical study of 24 teachers who received six-months of training to develop formative assessment practices, Wiliam, Lee, Harrison, and Black (2004) demonstrated the positive impact of using assessment for learning in classrooms.

However, the generalizability of these findings has been questioned by some. For instance, Bennett (2011) has argued that the mean effect size computed by Black and Wiliam (1998) is based on studies that are too diverse to be meaningfully combined, and that the meta-analysis conducted by Nyquist (2003) was too narrowly focused on the college-level population. Similarly, Dunn and Mulvenon (2009) have raised methodological concerns that may limit the conclusions drawn by Wiliam, Lee, Harrison,

and Black (2004). Because of these issues, the specific principles of assessment for learning should be further addressed.

1.3. The principles of effective assessment for learning

MacDonald (2007), while studying the implementation of formative assessment in classrooms, suggests that although the concept of formative assessment may be attractive to teachers, their usual criticism is that it is challenging to utilize in teaching practice consistently. Teachers describe the need of customizing teaching practice for individual students in large classes as well as helping the students who they deem to be more challenging to work with. Their biggest test, however, is to conform to the enormous requirements of standardized curriculum in restricted time periods (Wissick & Gardner, 2008). Trumpower, Filiz, and Sarwar (2014) discuss the needs of teachers using formative assessment. They succinctly lay out the criteria (referred to as 'elements' in this paper) for teachers to carry out effective formative assessment that it must (1) assess higher order knowledge, (2) identify students' strengths and weaknesses, (3) provide effective feedback, and (4) be user-friendly. Further analysis of literature indicates that another element should also be added to their existing scheme, specifically teachers' ability to track student learning over time. This will be discussed as the fifth element.

1.3.1. Higher order knowledge

Involving the students actively and cognitively in the learning process helps the teachers take interest in developing their understanding of the subject matter instead of simple information intake. Bransford, Brown, and Cocking (1999) and more recently, Hwang and Chang (2011) discuss at length the significance of this process, where students also learn to discern how their own and their peers' conceptual awareness of the ideas compare with the criteria and learning outcomes. This emphasis on the process of learning stresses the need to develop students' higher order knowledge. Studies show that students who focus on the learning process and not only passing the test accomplish better results and retain information longer (Wiliam, Lee, Harrison, & Black, 2004).

1.3.2. Students' strengths and weaknesses

For formative assessment to be effective, it must also be specific (Trumpower & Sarwar, 2010). A student's particular strengths and weaknesses must be identified in order for the teacher to help scaffold learning. The teacher and potentially the students themselves can improve learning based on the remedial instruction provided once the gap is identified between existing knowledge and the unit objectives. Teachers can also obtain further evidence of student learning using wide-ranging tasks, using a range of media, and applying new and existing skills in new situations (Yorke, 2003). This also provides students with an opportunity to explore new subjects from different perspectives and to strengthen their understanding and ability to transfer knowledge to different contexts. Needs based on different situations in the classroom and of students with different backgrounds can also be met with the use of varied methods of instruction. This enables teachers to select from a broad range of tools to help different students reach their learning goals.

1.3.3. Useful feedback

Wiliam, Lee, Harrison, and Black (2004) reported what has now become a popular idea that feedback is most effective when given in a timely manner. The authors conducted a study in which the results indicated that when teachers provided formative feedback within or between teaching units in their classrooms, students' progress rates were almost doubled, compared to the control classrooms, over the course of the year. It is also essential to note that feedback should not be provided rapidly since students should be given an opportunity to think and reflect about the concepts before seeking help from teacher or peer feedback. Rushton (2005) further indicates that learning gains tend to be bigger when feedback is focused on the task with specific suggestions for improvement rather than focusing on student's personality, even if it is positive. Effective feedback should refer to clear criteria, learning goals, and expectations for performance. It is also important to balance the amount of feedback based on the needs of the student.

1.3.4. User-friendly

Bransford, Brown, and Cocking (1999) reports that formative assessment is not popular among many teachers as they view it as a needless surcharge to their existing unmanageable workloads. Therefore, it can be concluded that formative assessment must be user friendly in order to succeed and the entire formative assessment process must be easy to use if it has to become a viable option in the classroom (Trumpower & Sarwar, 2010). There are different perspectives on the process being user-friendly, however. A teacher's perspective, although rewarding, could involve a tediously long process of providing remedial feedback to each student whereas students could reap substantial benefits from the assessment feedback and be motivated to use it only if they feel it is easy to use and access.

1.3.5. Tracking learning over time

After the identification of the learning goal, teachers assess the existing knowledge level of the student about the unit. Thereafter, progress is measured in comparison to that level toward the goal. Teacher's regular interaction with students is a key feature in classrooms using formative assessment where evidence is collected over several intervals about the growing understanding of the new concepts (Ballantyne, Hughes, & Mylonas, 2002). Teachers can utilize techniques like sophisticated questioning and extended dialogues to explore students' thinking to detect any misconceptions and to help them correct the conceptual errors. Köller (2001) found positive effects on student's deep learning in several experiments and field studies where teachers gathered new evidence of student progress and understanding at successive stages.

2.   Technology assisted assessment for learning in statistics education

Statistics educators face challenges in addressing their students' individual needs and providing students with specific feedback, because introductory statistics class size is always large as well as the students in this class have different academic backgrounds and various level of preparation, (Im & Yin, 2009). Therefore, Aliaga et al. (2005) reveal technology should be used for not only computing numbers but also for exploring conceptual ideas as well as boosting student learning.

To meet these needs in statistics education, a few computer applications have been developed such as Korelle, Stats-mIQ and Simulation Assisted Learning Statistics (SALS). To begin with, Koralle consists of worked examples which are methodically linked with problem-solving tasks. Students are first required to solve a problem, and then a worked example demonstrating the correct solution procedure is presented. Next, students can compare their own solutions to the example information (Krause, Stark, & Mandl, 2009). Therefore, while students are interacting with Koralle, they receive only corrective feedback on problem solution procedure. However, the software fails to automatically identify students' errors of solving statistical problems, so effective feedback for addressing these errors cannot be given. Second, the Stats-mIQ contains a series of multiple-choice quizzes. Each selected answer corresponds to extensive feedback. If the correct answer is selected, given feedback is corrective. Otherwise, students receive metacognitive feedback for addressing misconceptions and biases (Kleitman & Costa, 2014). Although students are given corrective and metacognitive feedback in the Stats-mIQ, whether students comprehend knowledge received through metacognitive feedback cannot be measured due to lack of providing new problems associated with the feedback. Finally, SALS consists of four phases: Externalization, Reflection, Construction and Application. In Externalization phase, students are posed a question whose context is related to daily life. Through this question, students realize their own ideas on the statistics concepts they are learning. In Reflection phase, students are asked to express their own ideas that were externalized in the Externalization phase. In Construction phase, students are required to interact with the Dynamically Linked Multiple Representations (DLMRs) via clear learning guides revealing the definition of key concepts and manipulating procedures. Therefore, students can identify the relationships among different representations and then construct their own concepts. Application phase is the last phase allowing students to solve two problem situations in different contexts and with different solution paths. As a result, in SALS let students both build their own concepts by the way of manipulating DLMRs and apply their conceptualization to new problems (Liu, 2010).

Even though these computer applications provides immediate feedback on students' performance and let students transfer their conceptualizations to new problems, none of these computer applications assesses students' prior knowledge. Several studies claim that identifying students' prior knowledge is crucial. For instance, Krause, Stark, and Mandl (2009) argue that students having little prior knowledge cannot effectively compare their own solutions with the worked examples in Koralle so they might not benefit from the standardized feedback. In addition, Leppink, Broers, Imbos, van der Vleuten, and Berger (2012) and Kalyuga (2009) claim that using prior knowledge for structuring and integrating new information can help students develop the latter ability. Such knowledge integration can be done through graphic organizers such as concept maps (Yin, 2012). Thus, ideal technology assisted assessment for learning tool in statistics education should explore each student's prior knowledge (strengths and weaknesses), provide immediate feedback on their weaknesses and then identify whether the provided feedback works for every students.

We have been considering these needs and developing a concept mapping website (conceptmapsforlearning.com) based on the aforementioned principles of effective formative assessment in our research lab (Filiz, Trumpower, & Atas, 2012). Within the website, students create concept maps on a particular subject and then receive individualized feedback and associated instructional material (e.g., videos, website links, examples, problems, etc.) based on a comparison of their concept map and a subject matter expert's map. After studying the feedback and instructional material, teachers can

track their students' progress by having them create revised concept maps (Filiz, Trumpower, & Vanapalli, 2014).

3. Brief presentation of the concept maps for learning website

This section explores how the topics of concept maps are chosen and how students interact with the concept maps.

3.1. Choosing a topic

A topic is chosen if either students face difficulties in explaining concept relations underlying a formula, or, students have misconceptions on this topic.

Of particular relevance to this presentation, Analysis of Variance (ANOVA) is chosen. Castro Sotos, Vanhoof, Van den Noortgate, and Onghena (2007) reveal that although students can manipulate and carry out calculations (ex. ANOVA) with statistical data, they might still have several misconceptions when results from inferential techniques are interpreted. For instance, students may only pay attention to p-value and omit information related to interpreting effect size (Broers, 2009). Even though few animated ANOVA tables illustrating a few of their properties are existed (Stirling, 2010; Van der Merwe & Wilkinson, 2011), the number of studies on the conceptualization of ANOVA calculations is limited (Van der Merwe, 2012). Thus, for identifying each student's strengths and weaknesses as well as providing feedback on their specific weaknesses via the CMfL website, the following concepts are selected: F-ratio, Number of groups, Between-groups mean square, Within-groups mean square, Sum of squares between groups and Degree of freedom.

Some might critique that the number of concepts in this set is fewer than it should be. For example, Cooke (1999) and Goldsmith, Johnson, and Acton (1991) claim that there should be at least12 concepts in the set, whereas Casas-García and Luengo-González (2013) argue that a set should include no more than 12 concepts to accurately assess a domain. Therefore, it appears that there is a consensus about using around 12 concepts to examine the rigor of the domain. Some studies, however, have used fewer than 12 concepts to assess a domain but stress the need of the core concepts being endorsed by domain experts. For instance, Casas-García and Luengo-González (2013) used a set of 11 concepts to generate students' knowledge maps about the concept of angles. After a sample of Primary (Grade 3P to 6P), Secondary (Grade 1E to 4E), Olympiad (2E), and Mathematics Undergraduate students completed pairwise ratings of the concepts, the representative knowledge map of each group was generated. Once they were compared to each other and similarity indexes between the knowledge maps were computed, a Kruskal-Wallis test was performed. Results showed that the similarity between the knowledge maps increased as age and experience of the students increased (H= 15.252, 7 d.f., p= 0.0329). Similarly, Boring (2005) identified 10 core concepts of an introductory psychology course in order to compare students' essay scores (g), which are given by the human graders (holistic and analytic) and the computerized grader, with Pathfinder similarity index (C). For the holistic graders, $g = 0.414C - 0.048$, where $R^2 = 0.001$ and $p > 0.05$. Thus, the numbers show that the degree of Pathfinder network similarity did not prove to be a sound enough indicator of the grade given by the holistic graders. It was also found that the relationship between the Pathfinder network similarities and the score awarded by the analytic graders was not significant ($R^2 = 0.012$ and $p > 0.05$). On the contrary, the degree of similarity between the computerized

grader's Pathfinder network and the student essay writer's network generated a comparatively good indicator of the grade awarded for that paper ($R^2 = 0.247$ and $p = 0.06$). These studies are examples of how researchers who used less than 12 concepts presented in their results that the generated knowledge maps were valid measures of students' comprehension.

Thus, we decide to generate an expert map of the selected concepts. To do so, three expert are asked to perform the pairwise ratings of these concepts by using a scale of 1(less related) to 5(more related) (see Fig. 1.). Afterwards, these ratings are converted to binary values through the PathFinder scaling algorithm (see Filiz, Trumpower, & Vanapalli, 2014). Then, the expert map (see Fig. 1.) is generated based these computed values. Also, the accurateness of this expert map is confirmed through the formula used for computing F-ratio.
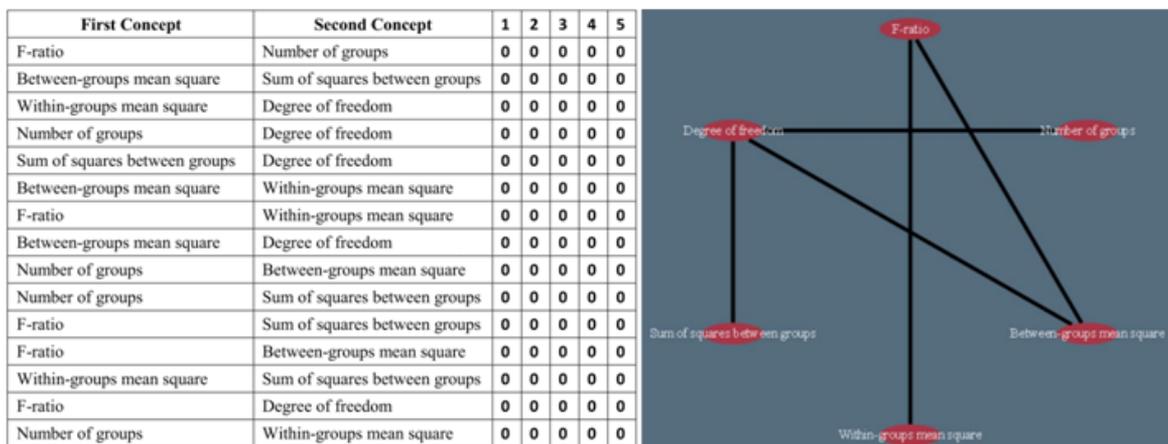
First Concept	Second Concept	1	2	3	4	5
F-ratio	Number of groups	0	0	0	0	0
Between-groups mean square	Sum of squares between groups	0	0	0	0	0
Within-groups mean square	Degree of freedom	0	0	0	0	0
Number of groups	Degree of freedom	0	0	0	0	0
Sum of squares between groups	Degree of freedom	0	0	0	0	0
Between-groups mean square	Within-groups mean square	0	0	0	0	0
F-ratio	Within-groups mean square	0	0	0	0	0
Between-groups mean square	Degree of freedom	0	0	0	0	0
Number of groups	Between-groups mean square	0	0	0	0	0
Number of groups	Sum of squares between groups	0	0	0	0	0
F-ratio	Sum of squares between groups	0	0	0	0	0
F-ratio	Between-groups mean square	0	0	0	0	0
Within-groups mean square	Sum of squares between groups	0	0	0	0	0
F-ratio	Degree of freedom	0	0	0	0	0
Number of groups	Within-groups mean square	0	0	0	0	0

Fig. 1. A pairwise rating task and the expert map of ANOVA

Thereafter, teachers using the website choose the topic "ANOVA" from amongst the topics. Students are then provided with the concepts corresponding to the chosen topic and rate the degree of relationship between the concepts in order to generate their concept map. The website then compares each student's concept map with the expert concept map to generate the individualized feedback for each student. This comparison is scored via C (configural similarity) measure which is computed by dividing the number of common links in two knowledge maps by the total number of links in both knowledge maps (Clariana, 2010). Through this similarity index, teachers find out degree of closeness between students maps and the expert map. Apparently, the value of this measure varies between "0" and "1". On the assumption that this value is always smaller than "1", our basic research question is whether or not various individualized feedbacks on the CMfL website help make this value much closer to "1".

In the following section, there are two examples of students' concept maps to demonstrate how students having different concept maps receive individualized feedback on their incomplete understanding of ANOVA calculations.

3.2. Demonstration of students' interaction with the website

After the rating task is performed, students receive three different types of feedback in the form of visual, textual, and linked instructional material.

First, visual feedback is comprised of a visual presentation of the expert concept map superimposed over the student's map with any discrepancies highlighted by different types of links. This visual feedback map consists of three different links *relevant link*, *extraneous link* and *missing links*.
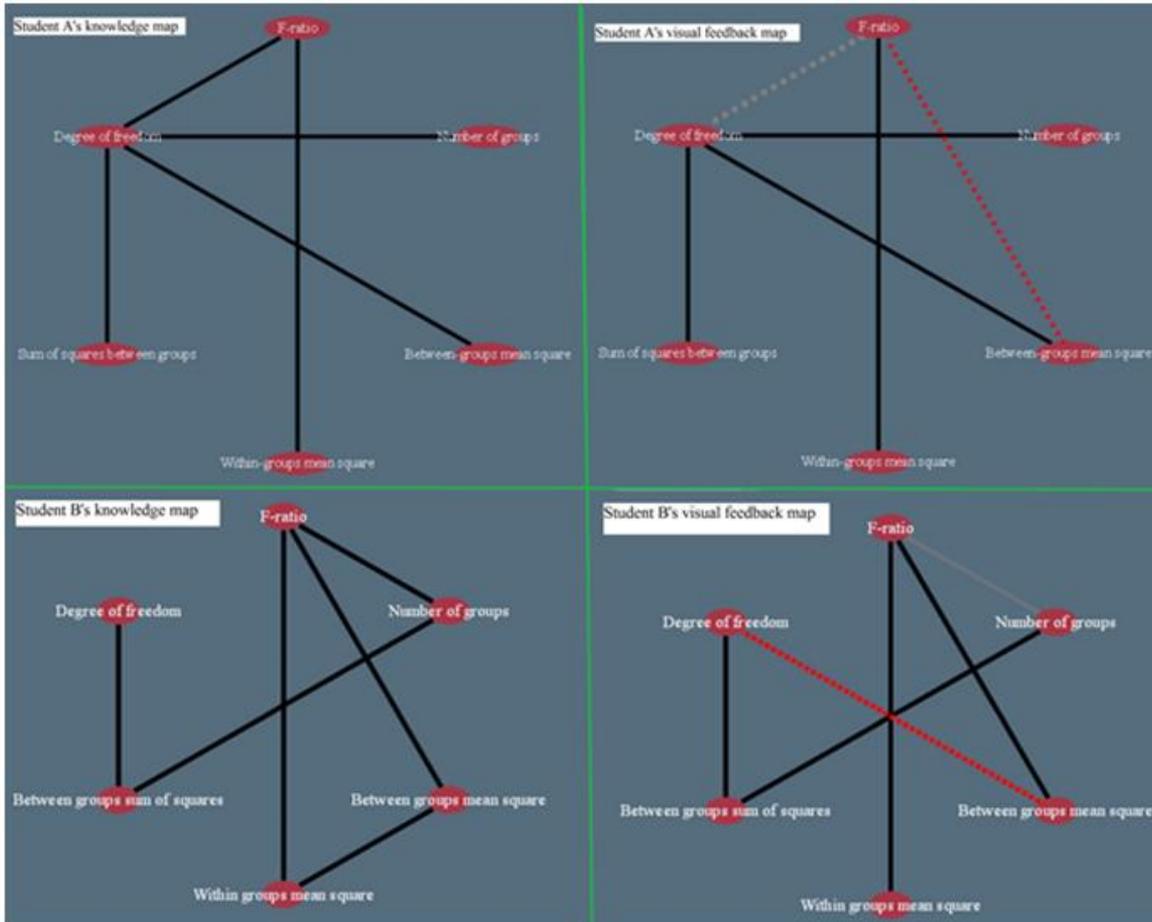

Fig. 2. An example of two students' visual feedback maps

As seen in Fig. 2, a black line appears provided that there is a link between two concepts in both an expert's map and the student map (ex. "*Degrees of Freedom* and *Number of Groups*" or "*Between Groups Degrees of Freedom* and *Number of Groups*"). This type of link will be referred to as a *relevant link*. A grey dotted line appears if there is a link between two concepts in the student map, but not in the expert's map (ex. "*Degrees of Freedom* and *F-ratio*" or "*Number of groups* and *F-ratio*"). This is referred to as an *extraneous link*. Finally, a red dashed line appears (ex. "*Between Group Means Square* and *F-ratio*" or "*Between Group Means Square* and *Degree of Freedom*") if there is a link between two concepts in the expert map, but there is no link between these concepts in the student map. Such links are referred to as *missing links*.

In addition to this visual feedback, additional instructional material is linked to any missing links. First, when students move the mouse cursor over a missing link, a text message appears which explains how the associated concepts are related; these

explanations have been provided by subject matter experts, but can be modified by individual teachers using the website. As seen in Fig. 3,, when Student A clicks on the missing link between "*Between Group Means Square*" and "*F-ratio*", a text message pops up as "*The F-ratio can be thought of as a measure of how different the means (ex. Between groups mean square) are relative to the variability within each sample*". Likewise, when Student B clicks on the missing link between "*Between Group Means Square*" and "*Degree of Freedom*", a text message appears which reads "*Once the sums of squares have been computed, between groups mean square is computed by dividing sums of squares by the degrees of freedom*". Second, if students double click on a missing link, they are able to access linked instructional material intended to illustrate the ways in which the associated concepts are related (e.g., videos, website links, examples, problems, etc.); again, this material has been provided by subject matter experts, but additional material can be added by individual teachers (see Appendix a).
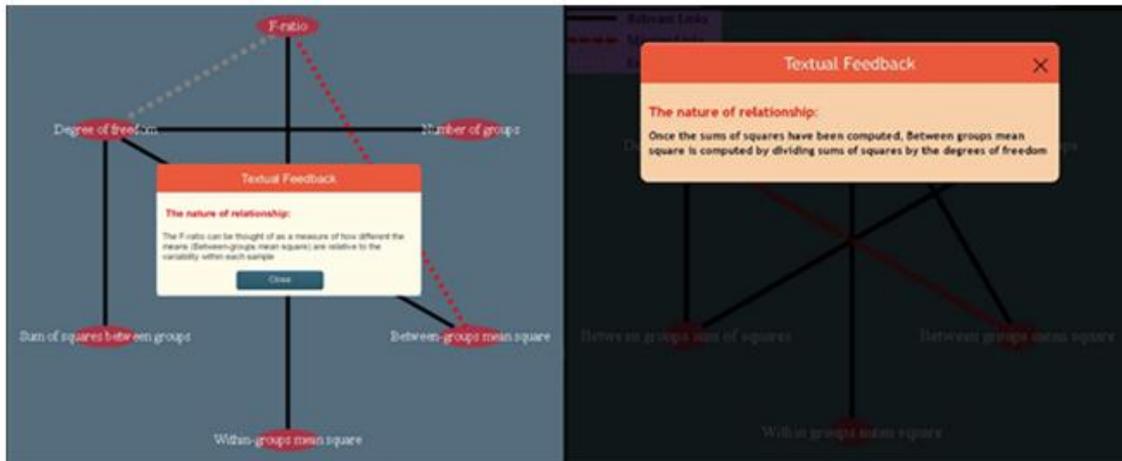

Fig. 3. An example of two students' textual feedback maps

4. How digital concept mapping website contributes to assessment for learning practices

One of the crucial contributions of the CMfL website to assessment for learning is to automatically identify each student's misunderstandings and misconceptions. Students taking introductory statistics courses are more like to have diverse backgrounds and prior knowledge. Consequently, students getting same total score from the exam cannot have similar strengths and weaknesses (Im & Yin, 2009). Therefore, identifying each strengths and weaknesses is crucial. Typically, the task of diagnosing misunderstandings is a difficult one, requiring both specific skills (e.g. how to conduct an interview) and plenty of time for developing an appropriate assessment task and for evaluating it (Browning & Lehman, 1988). Conversely, teachers are required to provide very minimal input in concept maps for learning website. Although they may add their own explanations and linked content, they are only required to choose a topic and submit a list of students for whom they wish to grant access to the website.

An additional contribution of the CMfL website is that it is capable of providing immediate feedback. Shute (2008) notes that although high-achieving students may benefit from delayed feedback, immediate feedback might be more useful for low-

achieving students. The author also reveals that immediate feedback is required for difficult tasks which are associated with higher order knowledge. Krause, Stark, and Mandl (2009) also note that lecturers save time via automatic feedback which provides immediate and individualized feedback in large size classes such as introductory statistics classes.

Moreover, providing feedback on the weaknesses of students in different forms is another of the crucial contributions of the CMfL website to assessment for learning, for research has shown that students are most likely to ignore verbal and written feedback (Shute, 2008; Lee, 2009). The concept maps for learning website provides both visual feedback and linked feedback as a form of associated learning activities including videos, games, or cartoon. Especially, students get informed about the similarities and differences between their maps with the expert map through visual feedback map. Krause, Stark, and Mandl (2009) find that students with little prior knowledge may ineffectively compare their own solutions with the worked example. Likewise, these students would not effectively compare their maps with the expert map if these maps were given side by side.

Furthermore, the concept maps for learning website is most likely to promote equitable education (Gikandi, Marrow, & Davis, 2011), because each student's weaknesses (misconceptions and/or misunderstanding) and strengths is diagnosed through computer based assessment for learning software. In addition, students are able to study received feedback from a variety of associated instructional multimedia materials. As a result, the CMfL website gives an opportunity to every student for studying on their misunderstandings as long as they need.

Finally, our concept maps for learning website may help students improve their general problem solving skills. In a related study, Schacter, Herl, Chung, Dennis, and O'Neil (1999) found that computer based concept mapping tasks improve students' problem solving performance. Particularly in this study, we demonstrate how students' ability of performing ANOVA test might be boosted by increasing their conceptual knowledge of this formal procedure through the CMfL website.

5.   Conclusion

According to Kleitman and Costa (2014), formative assessment tools consist of structured items such as multiple choice questions, mostly attached with structured instructions and extensive feedback. Since these tools supply timely and decisive diagnostic information, they should be used in introductory statistics classes. In general, these tools help students to complete gaps in their knowledge while addressing their certain misconceptions (Broers, 2009; Im & Yin, 2009). Unlike these formative assessment tools using multiple choice questions such as Koralle, Stats-mIQ and SALS, the assessment component of the CMfL website is concept maps. Therefore, this website has potential to offer further benefits.

One of the benefits of the CMfL website is that it can aid the instructors by automating many tasks in formative assessment. It also generates immediate and delayed feedback in variety of forms after automatically identifying each student's misunderstandings and misconceptions of the subject area. The CMfL website also has the potential to promote equitable education as well as helping students improve their general problem solving skills. This paper discusses whether these favorable contributions to assessment for learning are valid in statistics education.

For future studies, we are planning to create more concept maps related to different subjects. We will examine whether these promising contributions to assessment for learning are valid in different subjects.
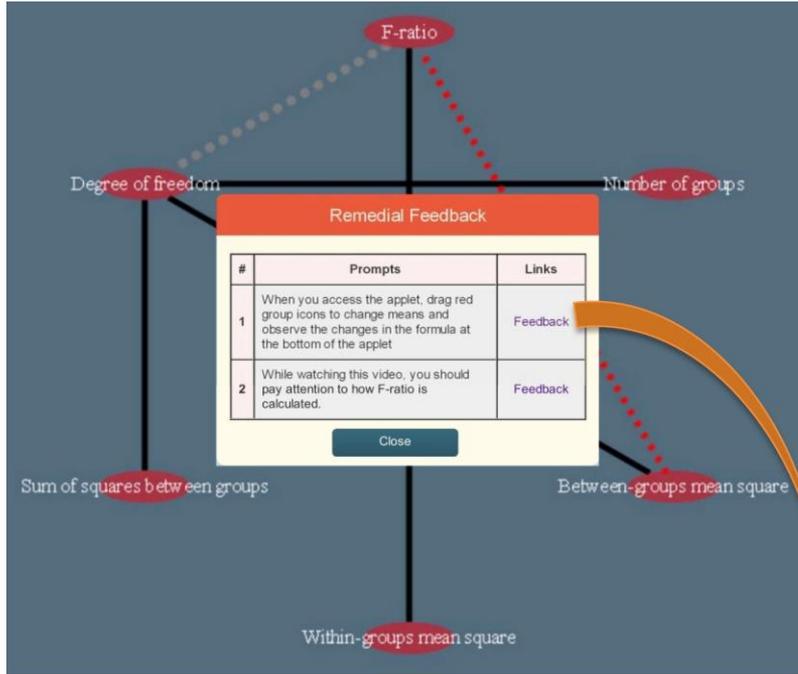
Acknowledgements

References

Aliaga, M., Cobb, G., Cuff, C., Garfield, J., Gould, R., Lock, R., Moore, T., Rossman, A., Stephenson, R., Utts, J., Velleman, P., & Witmer, J. (2005). *Guidelines for assessment and instruction in statistics education (GAISE): College report*. American Statistical Association. Alexandria VA. Retrieved from www.amstat.org/education/gaise

Ballantyne, R., Hughes, K., & Mylonas, A. (2002). Developing procedures for implementing peer assessment in large classes using an action research process. *Assessment & Evaluation in Higher Education, 27*(5), 427–441.

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice, 18*(1), 5–25.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7–74.

Black, P. J. (1993). Formative and summative assessment by teachers. *Studies in Science Education, 21*, 49–97.

Boring, R. L. (2005). The validity of human and computerized writing assessment. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 49, pp. 759–763). SAGE Publications.

Bransford, J. D., Brown, A., & Cocking, R. (1999). *How people learn: Mind, brain, experience, and school*. Washington, DC: National Academy Press.

Broers, N. J. (2009). Using propositions for the assessment of structural knowledge. *Journal of Statistics Education, 17*(2).

Browning, M. E., & Lehman, J. D. (1988). Identification of student misconceptions in genetics problem solving via computer program. *Journal of Research in Science Teaching, 25*(9), 747–761.

Casas-García, L. M., & Luengo-González, R. (2013). The study of the pupil's cognitive structure: the concept of angle. *European Journal of Psychology of Education, 28*(2), 373–398.

Castro Sotos, A. E., Vanhoof, S., Van den Noortgate, W., & Onghena, P. (2007). Students' misconceptions of statistical inference: A review of the empirical evidence from research on statistics education. *Educational Research Review, 2*(2), 98–113.

Cizek, G. J. (2010). An introduction to formative assessment: History, characteristics, and challenges. In H. Andrade & G. Cizek, (Eds.), *Handbook of formative assessment* (pp. 3–17). New York: Taylor and Francis.

Clariana, R. B. (2010). Deriving individual and group knowledge structure from network diagrams and from essays. In D. Ifenthaler, P. Pirnay-Dummer, & N. M. Seel (Eds.), *Computer-based diagnostics and systematic analysis of knowledge* (pp. 117–130). New York, NY: Springer.

Cooke, N. J. (1999). Knowledge elicitation. In *Handbook of applied cognition* (pp. 479–

510). Retrieved from http://www.cerici.org/documents/Publications/Durso%20chapter%20on%20KE.pdf

Dunn, K. E., & Mulvenon, S. W. (2004). A critical review of research on formative assessment: The limited scientific evidence of the impact of formative assessment in education. *Practical Assessment, Research & Evaluation, 14*(7), 1–11.

Filiz, M., Trumpower, D. L., & Atas, S. (2012). Analysis of how well a concept mapping website conforms to principles of effective assessment for learning. In A. J. Cañas, J. D. Novak, & J. Vanhear, (Eds.), *Concept Maps: Theory, Methodology, Technology - Proceedings of the 5th International Conference on Concept Mapping*.

Filiz, M., Trumpower, D. L., & Vanapalli, A. (2014). Exploring the mobile structural assessment tool: Concept maps for learning website. *Revista Colombiana de Estadística, 37*(2), 297–317.

Gikandi, J. W., Morrow, D., & Davis, N. E. (2011). Online formative assessment in higher education: A review of the literature. *Computers & Education, 57*, 2333–2351.

Goldsmith, T. E., Johnson, P. J., & Acton, W. H. (1991). Assessing structural knowledge. *Journal of Educational Psychology, 83*(1), 88–96.

Good, R. (2011). Formative use of assessment information: It's a process, so let's say what we mean. *Practical Assessment, Research and Evaluation, 16*(3), 1–6.

Greenstein, L. (2010). *What teachers really need to know about formative assessment*. Alexandria, VA: ASCD.

Hwang, G.-J., & Chang, H.-F. (2011). A formative assessment-based mobile learning approach to improving the learning attitudes and achievements of students. *Computers & Education, 56*(4), 1023–1031.

Im, S., & Yin, Y. (2009). Diagnosing skills of statistical hypothesis testing using the rule space method. *Studies in Educational Evaluation, 35*(4), 193–199.

Irons, A. (2008). *Enhancing learning through formative assessment and feedback*. Taylor & Francis.

Kalyuga, S. (2009). Instructional design for the development of transferable knowledge and skills: A cognitive load perspective. *Computers in Human Behavior, 25*(2), 332–338.

Kleitman, S., & Costa, D. S. J. (2014). The role of a novel formative assessment tool (Stats-mIQ) and individual differences in real-life academic performance. *Learning and Individual Differences, 29*, 150–161.

Köller, O. (2001). Mathematical world views and achievement in advanced mathematics: Findings from the TIMSS population 3. *Studies in Educational Evaluation, 27*, 65–78.

Krause, U.-M., Stark, R., & Mandl, H. (2009). The effects of cooperative learning and feedback on e-learning in statistics. *Learning and Instruction, 19*(2), 158–170.

Lee, I. (2009). Ten mismatches between teachers' beliefs and written feedback practice. *ELT journal, 63*(1), 13–22.

Leppink, J., Broers, N. J., Imbos, T., van der Vleuten, C. P., & Berger, M. P. (2012). Self-explanation in the domain of statistics: an expertise reversal effect. *Higher Education, 63*(6), 771–785.

Liu, T.-C. (2010). Developing simulation-based computer assisted learning to correct students' statistical misconceptions based on cognitive conflict theory, using 'Correlation' as an example. *Educational Technology & Society, 13*(2), 180–192.

MacDonald, M. (2007). Toward formative assessment: The use of pedagogical documentation in early elementary classrooms. *Early Childhood Research Quarterly, 22*(2), 232–242.

Nyquist, J. B. (2003). *The benefits of reconstruing feedback as a larger system of formative assessment: A metaanalysis*. Unpublished master's thesis. Nashville, TN, Vanderbilt University

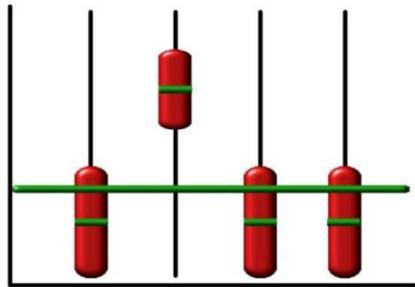Rushton, A. (2005). Formative assessment: A key to deep learning? *Medical Teacher,*

27(6), 509–513.

Schacter, J., Herl, H. E., Chung, G. K. W. K., Dennis, R. A., & O'Neil, H. F. Jr. (1999). Computer-based performance assessments: A solution to the narrow measurement and reporting of problem-solving. *Computers in Human Behavior, 15*, 403–418.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research 78*(1), 153–189.

Stirling, D. (2010, July). Improving lectures with CAST applets. In C. Reading (ed.), *Data and context in statistics education: Towards an evidence-based society. - Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8).*

Trumpower, D. L., Filiz, M., & Sarwar, G. S. (2014). Assessment for learning using digital knowledge maps. In D. Ifenthaler & R. Hanewald (Eds.), *Digital Knowledge Maps in Education* (pp. 221–237). Springer New York.

Trumpower, D. L., & Sarwar, G. S. (2010). Effectiveness of structural feedback provided by Pathfinder networks. *Journal of Educational Computing Research, 43*(1), 7–24.

Van der Merwe, L. (2012). *An analysis of the field of statistics education research with special reference to computer-assisted teaching and learning*. University of the Free State. Retrieved from http://etd.uovs.ac.za/ETD-db/theses/available/etd-07182013-103342/unrestricted/VanDerMerweL.pdf

Van der Merwe, L., & Wilkinson, A. (2011). Mapping the field of statistics education research in search of scholarship. *International Journal for the Scholarship of Teaching and Learning, 5*(1), 29.

Wiliam, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education: Principles, Policy, & Practice, 11*(1), 49–65.

Wissick, C. A., & Gardner, J. E. (2008). Conducting assessment in technology needs: From assessment to implementation. *Assessment for Effective Intervention, 33*, 78–93.

Yin, Y. (2012). Using tree diagrams as an assessment tool in statistics education. *Educational Assessment, 17*(1), 22–49.

Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education, 45*(4), 477–501.
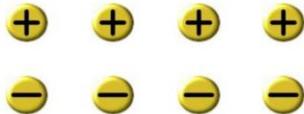
Appendix a