Evaluating Metrics for Automatic Mind Map Assessment in Various Classes

Amber Franklin
Department of Speech Pathology and Audiology
Miami University
Oxford, Ohio 45056
Email: franklad@miamioh.edu

Peter Jamieson
Department of Electrical and Computer Engineering
Miami University
Oxford, Ohio 45056
Email: jamiespa@miamioh.edu

Walter Vanderbush
Department of Political Science
Miami University
Oxford, Ohio 45056
Email: vanderw@miamioh.edu

Tuo Li
Department of Electrical and Computer Engineering
Miami University
Oxford, Ohio 45056
Email: lit8@miamioh.edu

Julie Semlak
Department of Media, Journalism, and Film
Miami University
Oxford, Ohio 45056
Email: semlakj@miamioh.edu

***Abstract*—Over the past three years, we have been studying how automated evaluation of student mind maps (when compared to an expert map) shows student learning for a variety of metrics. The goal of this work is to build a system that would then allow students to evaluate their understanding of the terminology in a respective field. The weakness of our studies, so far, is that our focus group to study these metrics consists of a single course in the field of Computer Engineering, and though this class has been used over multiple years to demonstrate the feasibility of our approach in a longitudinal study, a broader study needs to be done. In this work, we show how our current metrics perform across three additional fields; specifically, we have collaborated with instructors in speech pathology, communications, and political science (in addition to our traditional class in computer engineering). We then use our methodology to determine if these courses and a term long mind map exercise have similar results than previously reported and are these results evidence of student learning. Our results show that our existing metrics have similar results for one of the three new courses. However, in the two other courses, the data shows no evidence of learning based on the mind map exercise. Each of the instructors of these courses describes their experience with the activity. Additionally, we evaluate the construction of the expert maps in each course to understand if there is a graph-based structural reason why we the results might be different. We conclude by suggesting our methodology is good for courses where terminology is clearly defined and is used and studied throughout the semester, and describe some future directions for this research.**

I. Introduction

The overall goal of this research is to investigate if machines can help provide students with feedback on their learning using mind maps. Mind maps are simple visual drawings that includes terms of interest (which we will call nodes) and lines connecting these terms (which we will call edges). We use the terms nodes and edges as they are common terms in for describing components of a graph, which is a structure under study in a mathematical field called graph theory. Based on this, we are examining if graph theory and related algorithms can help us evaluate a student's mind map. This evaluation is achieved by comparing a criterion map to the students map, where the criterion map is the original mind map created by an expert (instructor). The criterion map determines the number of words and their connectivity.

To achieve this, we have proposed and studied a number of metrics in previous publications ([1], [2], and [3]) in terms of finding a set of metrics that show if a student is learning. The weakness in these works is that they have focused on a single course (Digital Systems Design) within a single major (Computer Engineering), and even though the results for some of our metrics (two in particular) have shown promising trends over each year of data collection, the bigger question is if our methodology shows similar trends over a range of courses.

The goal of this paper, then, is to present a broader range of courses and apply the same experimental method and evaluation to see how these metrics show learning. Specifically, we have enlisted the help of three instructors in speech pathology, communication, and political science and implemented our experimental method on some of their courses. The data collected over a semester in each of these courses was then analyzed using two of our most promising metrics, and the results are evaluated in this paper. In particular, we found that our favorite metric (a match metric) performed well in one case (speech pathology), showed some indication of learning in a second class (political science), and showed little evidence of learning in the third class (communication).

After presenting the methodology, a brief introduction to the metrics, and the results, we provide each instructors perspective on experiment, mind map activity, and their respective results. This discussion provides us with a conversational explanation for our results, and additionally, we look at aspects of the criterion maps to see if there is any difference that can further explain why in two out of four cases our preferred metric seems to be capturing student learning. This work provides us with one recommendation that our methodology seems to be applicable to courses that introduce terminology with clear definitions and that terms are used throughout the semester. For courses that do not fit this model, we suggest some future work that might allow our techniques to be useful to students in their learning process.

The remainder of this paper is organized as follows: section II provides a background on mind maps and research into their use as measurement artifacts. Section III describes our methodology and measurement metrics. Section IV presents the results for two of our measurement metrics for the 4 courses we used in 2012/13, and describes these results in some details. Section V provides the three new course bases with a more personal description of their experience with the experiment and their respective results. Section VI discusses our overall results and concludes the paper.

II. Mind Maps as Measurable Artifacts

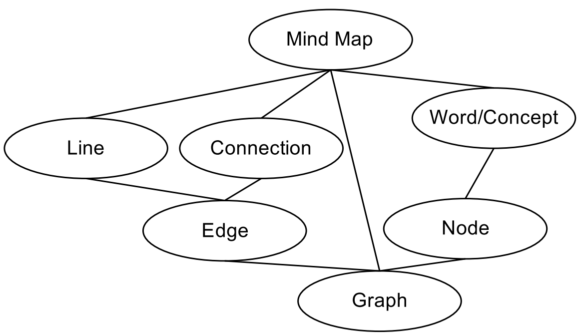

Fig. 1. Example of a mind map on the relationship between mind maps and graphs

Mind maps are simple visual representations of terminology and simple one-to-one relationships [4]. Mind maps [5] are a visual representation that are used in a number of settings including a Class Assessment Techniques (CATs) [6] which allows teachers to evaluate student understanding in class and provide feedback. Figure 1 shows an example mind map that expresses the main author's understanding of mind maps and how they relate to mathematical graphs. The words/concepts that are in a mind map are the nodes of a graph (circled bubbles), and the connecting lines between these words are edges of a graph.

There is continued interest in mind maps as a pedagogical tool that can help structure learning [7] as well as a vast and rich data set that can provide other insights [8]. Mind mapping tools are readily available and the technique is used in an wide number of areas. In our past studies we have provided background on mind maps in scoring. An updated one for mind maps includes:

- comparing the scores on tests to the technique [9]
- having two independent experts score (sometimes with a rubric) the mind map on a scale two times with one week delay and compare correlation of ratings [10]
- using structures and frameworks to identify redundancies and troubling portion of a map [11]
- using a large data-set of mind maps for deep fact finding of interrelated topics [12]

The type of mind maps we use in this study are called closed, which means they have a limited set of words (nodes) [13]. Our scoring technique that uses criterion maps is called *comparison with a criterion map* [14].

III. Experiment and Measurement Metrics

To perform automatic feedback from mind maps to help students learn a field's vocabulary, our focus has been on semester long experiments that evaluate student mind maps with various metrics. In this work, we continue this experimental setup, but for a broader range of courses. In this section, we will describe the experimental setup and the measurement metrics that we have found to be most useful in showing that students have learned.

A. Experimental Method

We start with the assumption that *most* students learn about a particular area (course) over a semester, and this learning includes a better understanding of the technical vocabulary and terms as related to a class topic/field. Nation [15], Coxhead [16], and Chung *et. al.* [17] discuss the relevance of technical vocabulary in a field and quantify that their are approximately 5% technical, field-based the words used in related publications. However, the technical vocabulary is not necessarily the most important learning objective, and in terms of emphasis and assessment, the vocabulary might be a periphery outcome.

To measure the relevance of the vocabulary we will use Wiggins and McTighe's simple taxonomy [18]:

1) worth being familiar with
2) important to know and do
3) enduring understanding

where the importance of the learning objective is more important as the number increases. Each of our instructors in this experiment will give their courses technical vocabulary a 1, 2, or 3 rating depending on how important it is to their respective course and students.

The experiment is a semester based longitudinal experiment. First, the instructor picks twenty terms (closed mind map) that they then use to create the criterion map (expert map). Next, the students are briefly taught how to make a mind map using an example that is not part of their class. For example, a list of countries is provided and the instructor shows how different mind maps could be created depending on the context; for example, countries organized by trading partners versus geographical location versus immigration paths will each make a very different mind map. Then over the semester the students will be given the 20 terms (unorganized) and will create an in class mind map.

In terms of controls, we have attempted to control the following:

- Each in class activity is kept to 10 minutes
- There are twenty terms to create a mind map and those terms stay the same each time
- The activity is done at least twice, but preferably three or more times spread somewhat evenly throughout the semester

Students create their mind maps on paper, and therefore, the mind maps need to be electronically encoded to allow the measurement metrics to be calculated. This process takes some time, but until we have a flexible electronic tool this is part of our methodology. The conversion of paper mind maps is done once the semester is completed and grades have been submitted. Research consent forms are opened and students that gave permission for their mind maps to be included in the experiment are kept and non-participants mind maps are shredded and removed from the data set.

Once the criterion maps and the student mind maps have been converted into an electronic format, we run our data analysis tools to generate the respective metrics as will be presented in the results section.

B. Metrics to Measure Mind Maps

In our past papers ([1], [2], and [3]) our goal has been examining different metrics that measure how similar the criterion map is to a student mind map. Over these studies, we have found two metrics that seem to strongly show differences and a third one that is pretty good at showing graph similarities. In this work, we will use one of the strong metrics (the second metric was being evaluated) and the okay metric; these metrics are called *match metric* and *RGF-distance*, which we will describe below. These two metrics are the best performing metrics in our 2013 paper [2], which has additional description of why these metrics seem to better capture similarities and differences as related to learning.

The *match metric* is an edge by edge comparison between the student mind map and the criterion map. The nodes in our graphs are uniquely identified by a label (the term written in the bubble), and this allows us to compare the two graphs in linear time. During this comparison a number of statistics are recorded about the differences including missing nodes ($MissN$), extra edges ($ExtraE$), and matching edges ($MatchE$) where the comparison is the student map as compared to the criterion map. The *match metric* is a combination of these statistics:

$$MatchMetric = \frac{MatchE}{MissN + ExtraE + MatchE} \qquad (1)$$

This equation results in a number between 0 and 1. The number is interpreted where as it approaches 1 indicates there is more similarity between the two graphs.

The second metric is *RGF-distance*. This metric is a little more complicated and relies on what are called, graphlets, which, are “a connected network with a small number of nodes” [19] and these small graphs are non-isomorphic induced subgraphs of a larger graph. Figure 2 shows all the graphlets of size 2, 3, and 4.

The existence of graphlets is used to analyze the structure of a graph. The procedure developed by Przulj *et. al.* [19] is to search for all graphlets of size 3, 4, and 5 in a given graph. Based on the count of each type of graphlet, a signature is constructed in the form (g1, g2, g3, g4, g5, g6, g7, g8, ..., g28, g29), where g1 is number of the first type of graphlet of size 3 shown in figure 2 and g29 is the count for the last graphlet of size 5. This signature can be compared to another graphs’ signature to get a measure of similarity, and Przulj *et. al.* used their technique to compare graphs representing biological structures such as proteins.

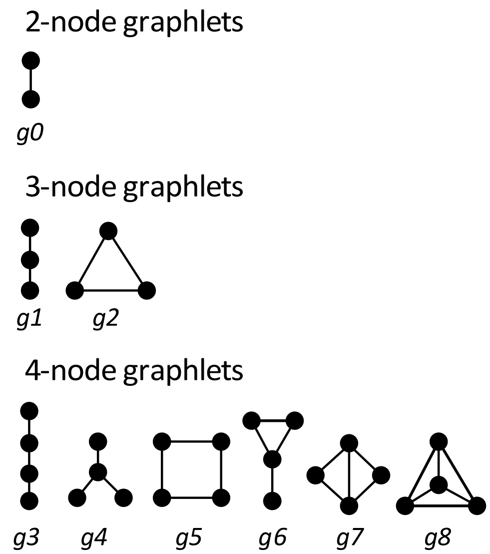

Fig. 2. Graphlets of size 2, 3, and 4

RGF-distance is a measure of the difference in frequency of graphlets of g1, g2, g3, ..., g28, and g29 appearing in the two graphs being compared. A tool called GraphCrunch II will calculate this metric, and as this value approaches 0 the more similar the two compared graphs are, however, this is an approximate measure since the specific labels of the nodes are not used.

IV. Results

In this section, we show the results for all four courses (computer engineering, communication, speech pathology, and political science) for both metrics (Match Metric and RGF-distance). Additionally, we provide a table that summarizes some of the differences in each courses experimentation, some parameters of each criterion map, and a percentage of metric results that align with evidence of learning. Our key assumption in this exploration is that we assume that on average the students are learning the technical vocabulary in the course as the course proceeds and if a metric is capturing this in the mind maps we should see the majority of the the students metrics trend a certain way (towards 1 from 0 for the *match metric* and towards 1 from larger numbers for the *RGF-distance*).

Table I includes a summary of a number of pieces of data for our experiments. Column 1 shows the course area. Columns 2 and 3 report the percentage of students included in the experiment that have a measured metric (RGF-distance and Match Metric, respectively) that trends better in a student’s final mind map as compared to their initial mind map (for example, the match metric would be closer to 1 in an upward trend and the RGF-distance metric would be closer to 0 in a downward trend). Column 4 shows the number of students that gave permission to be included in the experiment and participated in all the mind map activities, and column 5 shows how many mind map activities were performed for the

TABLE I. DETAILS FOR EACH COURSE EXPERIMENT

	Metric Summary		Experimental Parameters		Criterion Map	
Course Area	Percent improved Match Metric	Percent improved RGF-distance	Population	Mind Map Activities	Average Degree	Density
Computer Engineering	100%	69%	32	3	2.85	0.30
Communication	27%	73%	51	4	0.95	0.10
Speech Pathology	91%	59%	34	3	1.68	0.18
Political Science	55%	57%	42	2	2.95	0.31

respective course. Finally, columns 6 and 7 report the average degree and density of the criterion mind map and these graph parameters will be used to describe differences in the criterion map.

Figure 3 shows each of the four courses and includes students metric if they participated in all the mind map activities. In terms of our assumption of learning, we hope to see an upward trend (from values of 0 towards values of 1) for the match metric. We see this upward trend in two cases - computer engineering and communication. Since there are many students in each of these courses, the data is a little noisy and is hard to interpret, but the summary of the "Percent Improved Match Metric" in Table I suggests that these 2 upward trending courses are clearly showing student learning. Conversely, both the graph and the summary data show that the communication course is not trending upwardly. The political science course is not showing clear results, and unfortunately since only two data points (mind map activities) were performed it is hard to make strong conclusions from this result.

Figure 4 shows each of the four courses and includes students metric if they participated in all the mind map activities. In terms of our assumption of learning, we hope to see a downward trend (from values tending towards 1) for the RGF-distance metric. The results here are much less pronounced compared to the match metric, and the summary data in Table I for "Percent improved RGF-distance" shows that all courses have over 50% of the students demonstrating improvement. On closer examination of the results, it is clear that RGF-distance and its estimation of graph similarity is a poor comparison metric since there can be false positives. For this reason we used these results to look at creating a new metric that mixes graphlets with the match metric in our paper [3].

We performed additional analysis on this data set including evaluating our previous experimental graph metrics and checking if there is any correlation between grades and the metrics. As reported previously with computer engineering and with these new courses, we saw little correlation between grades and any of our measurement metrics. No other metrics seemed to show trending as match metric and RGF-distance metric have.

V. INSTRUCTOR EXPERIENCES AND DISCUSSION

In this section, each of the three instructors (additional courses beyond our traditional computer engineering course) provide their experiences with the mind map activity and the respective results for their class.

A. *Amber Franklin - SPA 334 Clinical Phonetics and Articulation Disorders*

Clinical Phonetics and Articulation Disorders (SPA 334) is a required junior level class for students majoring in the department of Speech Pathology and Audiology. The majority of majors in this department plan to apply to graduate schools where they will be trained as Speech-Language Pathologists or Audiologist. The SPA 334 course addresses the fundamental aspects of phonetic transcription and articulatory phonetics. Students are taught to transcribe spoken English using the International Phonetic Alphabet. Other topics include English dialectal variation, and clinical phonetics as it applies to normal speech development and articulation disorders. The course quizzes and exams assess knowledge of terms and concepts as well as phonetic transcription ability, which is an applied skill. Some students excel in conceptual knowledge but do not perform well in transcription tasks. The mind map exercise corresponds to the conceptual and technical information taught but does not address the applied skill of phonetic transcription, which constitutes a significant portion of students' final grades. This distinction may partially explain why there was no correlation between mind map performance and final grades. The terminology as a learning outcome is rated as "2 -important to know and do" based on the taxonomy discussed earlier.

The SPA 443 course builds gradually on a foundation of terms and concepts. Each quiz in the course is cumulative, thereby encouraging students to retain information introduced earlier in the semester. Twelve of the 20 concepts on the list were introduced before the midterm (2nd data point) and the remaining eight concepts were introduced after the midterm. Additionally, several of the terms introduced before the midterm were revisited in the second half of the semester and linked explicitly to new concepts. It is likely that the cumulative nature of the course was reflected in the improved match metric and decreased RGF distance over time. The students understood that their maps were going to be compared to my criterion map once the course was done. However, they also knew that mind map performance would not affect their course grade, making this a very low-stakes activity. The students demonstrated focused attention when generating mind maps at each time point in the semester. I believe the low-stakes nature of the activity facilitated student engagement and allowed them to interact with the material in an authentic way rather than worrying about the "right answers." The technical vocabulary in this course is rated as "2-important to know and do" based on section III-A.

I observed several differences in the visual representations of student maps. One map was arranged in two distinct columns. Terms within a column were connected using curved

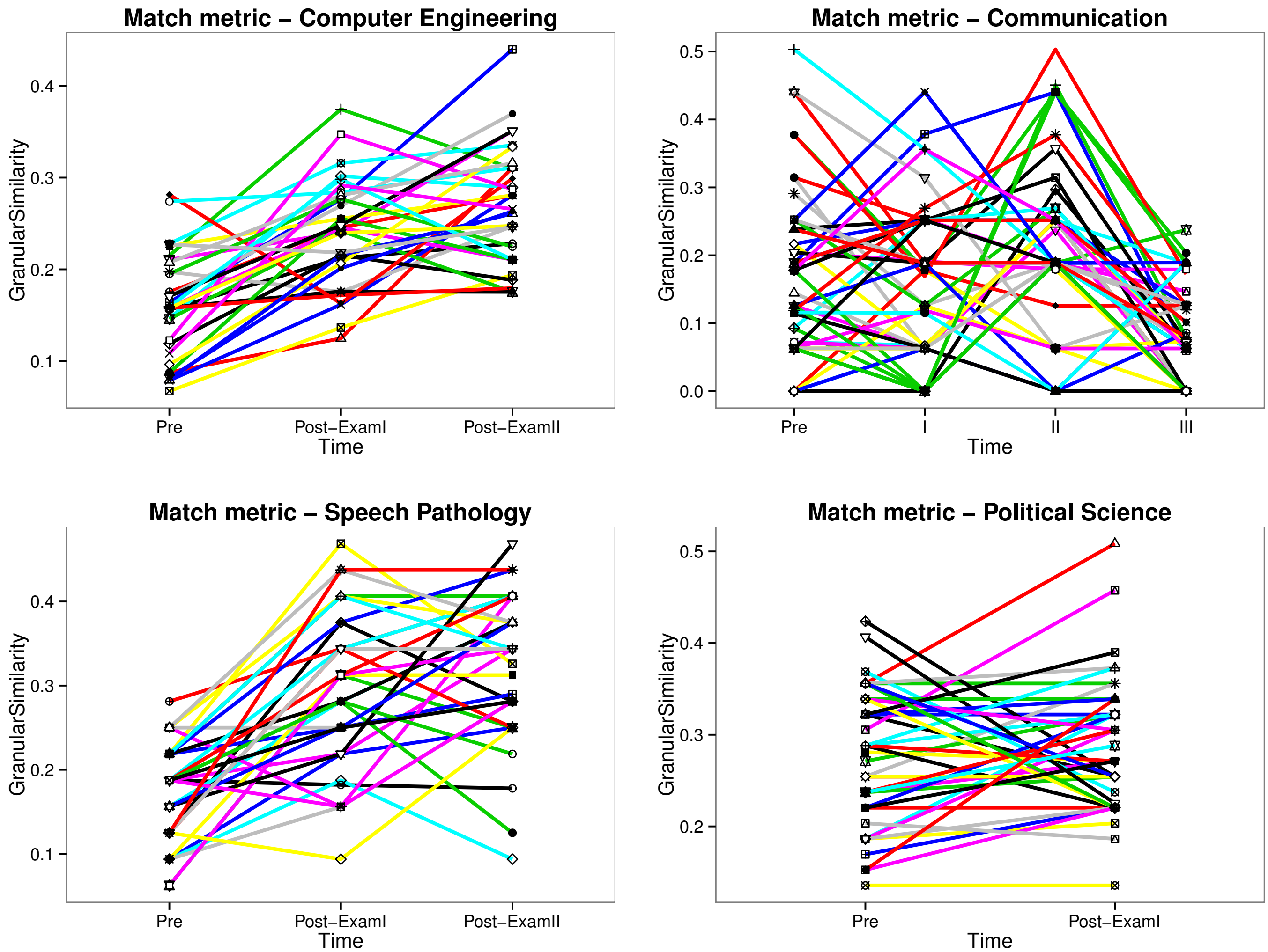

Fig. 3. Match Metric results for each of the four courses

lines and terms between columns were connected with straight lines. Other maps demonstrated a hierarchical structure, with a key concept at the top of the page and related concepts branching below it. Several other maps had a central concept in the center of the page with related terms and concepts branching like spokes from the center. Students were given a general example of a mind map at the beginning of the semester. However, they were not given specific directions regarding the visual structure of the map (e.g. "Start with a key concept in the middle of the page"). The electronic encoding of the maps removes the visual texture from the students' paper representations. Though not within the scope of the present study, I found it interesting to consider how the variation in written map representation may further reflect students' conceptual knowledge.

B. Walter Vanderbush - POL 101 Politics and National Issues

Political Science 101 is a course that fulfills a Liberal Arts social science requirement at Miami but does not count toward the political science major or minor. The primary goal of the course then is not to prepare students with the conceptual knowledge and analytical skills necessary in upper level political science classes, but rather to expose those who major in Engineering, Accounting, or Zoology to social scientific analysis of political issues and debates. In this version of the course, just over half of the students were in their first year of college; only one of the students was a political science major, and one other student was majoring in public administration, which is housed within the political science department.

Before the first map exercise, I did an exercise on the board with a list of countries and attributes, suggesting various ways that one might link them depending on the context. The concepts chosen for the mind map assessment range from democracy, capitalism, and freedom (which all students will have some familiarity with on day one of the semester) to pluralism, libertarianism, and judicial review (which students may have heard before but were not universally likely to fully understand) and a couple of concepts that only students following politics pretty closely were likely to have familiarity with (self-deportation, e.g.). Over the course of the semester, the expectation was that the relatively unfamiliar concepts would become familiar. On the other hand, the definitions of the familiar ones such as democracy and freedom are

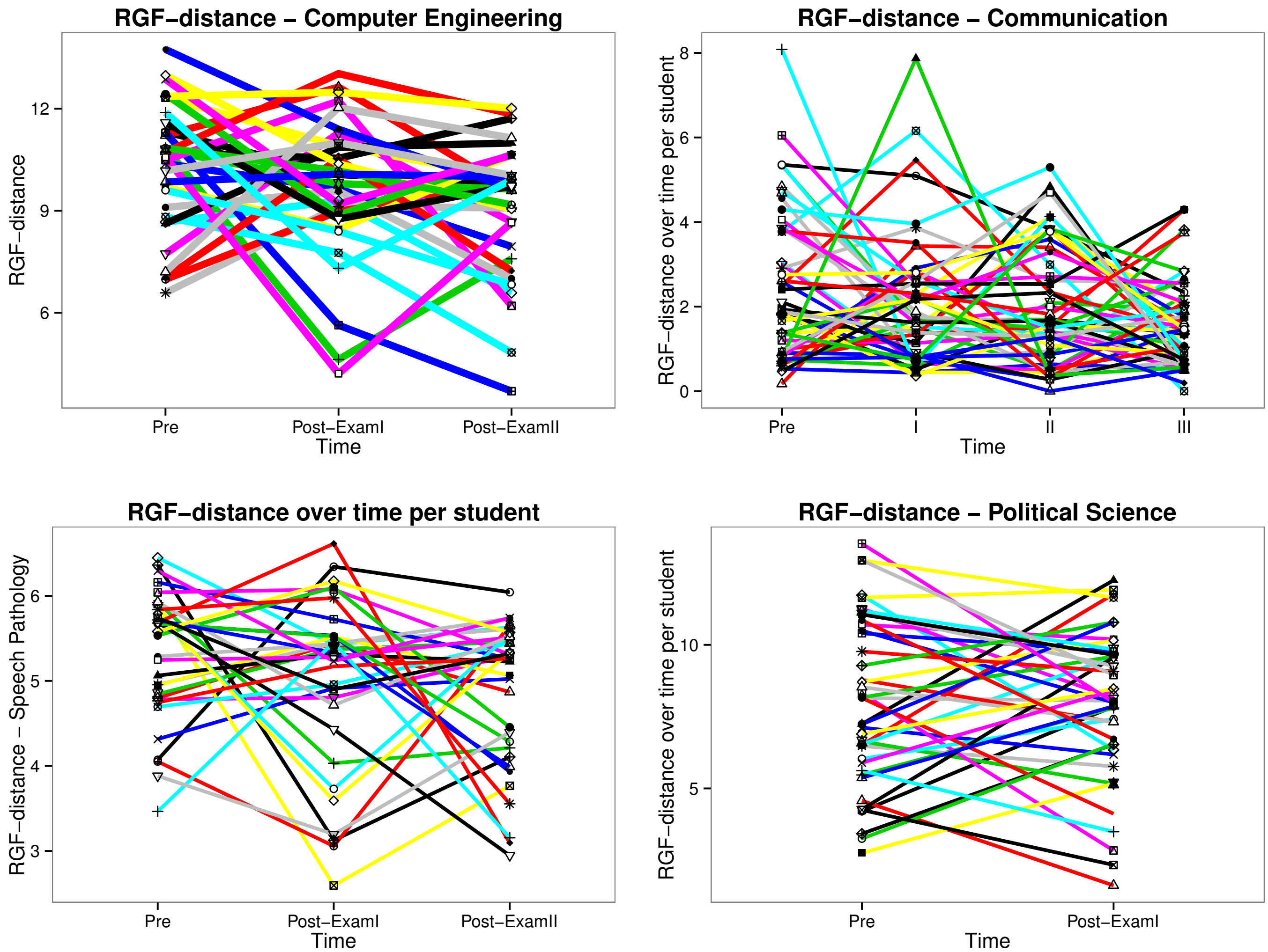

Fig. 4. Match Metric results for each of the four courses

challenged throughout the semester, as students are asked to think of them as contested concepts. To take another concept, separation of powers, one lesson students should learn during the semester is that some democracies have explicit separation of powers such as are laid out in the US Constitution, but other democracies do not see that sort of separation as necessary. On his or her mind map, a student is likely to have made a clear link between separation of powers and democracy at the beginning of the semester, thinking that all democracies had the separation of powers like the U.S., but by the end of the class might reasonably still make that link or might decide not to. The terminology as a learning outcome is rated as “2 -important to know and do” based on the taxonomy discussed in section III-A.

I expect that the mix of concepts contributed to the mixed results for the metrics in this study. I tried to make the list reflect the variety of discussions that the class would have during the course of the semester. To that end, there were big philosophical concepts such as democracy, freedom, and individualism; traditional political science concepts such as pluralism, separation of powers, and judicial review; and reflecting some of the contemporary policy debates, terrorism, American exceptionalism, and even self-deportation. The second of those groups is the closest to the idea of a technical vocabulary that made up a larger percentage of lists in other classes studied here. In a course intended to prepare students for more advanced political science classes, that technical vocabulary would have been more central. This initial use of mind maps, as well as pre-course discussions with my colleagues did lead me to think more about the ways that my twenty concepts might be interrelated. My criteria map tended to make more linkages than those of nearly all of my students. Many of them seemed to look for 2 or 3 strong relationships before moving on to the next concept. In my ”expert” map, several concepts had five or six connecting lines, and others even more.

C. Julie Semlak - COM 135 Introduction to Public Expression and Critical Inquiry

The communication course used for this study, an introduction to communication theories course, is a survey course intended to provide communication majors with a foundation of metatheoretical issues, as well as a survey of communication theories they may encounter throughout their studies. The terms used for this study were metatheoretical terms exploring the philosophy of theory development. Although these terms are utilized and reinforced throughout the semester, they are emphasized much more at the beginning of the course, and were likely most salient to students Post-Exam One, as these terms are the primary focus of exam one. Although the terms do appear on exam two and three, the content emphasized for these exams is specific to the individual theories taught during the semester.

Before students completed the mind map activity for this project, I introduced the definition of mind maps, and as a class we generated a mind map on the board, using pizza as the topic, and the word at the center of my sample mind map. I think this example prompted many of my student to want to put one of the concepts they had to work with at the center of their mind maps, which lead to some frustration when I told them they could choose any of the 20 words for the center. Similar to Dr. Franklin, although I also emphasized this activity was not graded, many students were concerned about completing their map correctly, and asked me if their mind map was correct. Also, after the second data point, several students asked to compare their second map to their first map. Those who did compare remarked at how different their respective maps were, and commented on the errors they had made in their data-point one map. The technical vocabulary in this course is rated as "1-worth being familiar with" based on the ratings in III-A.

When using mind maps as a pedagogical tool, one should take into account the overall goals of the course. The communication course used for this study's goals focus on students being introduced to a variety of communication theories, students being able to apply individual concepts or theorems from individual theories to their experiences, and to identify the advantages and disadvantages of specific theories. While these skills are important for effective communication, they are skills, which are often assessed different from knowledge or understanding. When considering mind maps as a pedagogical tool, the overall intent of the course should be considered, as mind maps may not be the best assessment tool for a skills-based course.

VI. Discussion and Conclusion

From our results and discussions we find that our study, which expanded the number of classes studied with automated mind map evaluation, shows that there are more details in achieving a successful automated feedback metric(s) for students. In particular, from this study we can see how the nature of the ideas, words, and concepts used within the CAT has significant impact on the quality of the results. Second, this study provides further evidence that our best metric *match metric*, because of its more direct comparative feature, is still the best metric we have to analyze students maps and provide detailed feedback of what they are and aren't connecting as compared to the expert.

Our mixed results on the *match-metric* groups the four courses into two sets. In one set, which includes computer engineering and speech pathology, the chosen words for the CAT have very clear definitions and relationships to other words, and these terms are introduced and used throughout the class. Conversely, in the other set, which includes political science and communication, the words used either do not have clear definitions that are to be developed over time, or the terms themselves are introduced early and are used sparingly throughout the term. It makes sense that the first set when analyzed by our metrics shows student improvement because the terms are easier to classify and continually used over the semester, which follows the experimental design. Therefore, if courses have terms that are clearly defined and consistently used then this methodology seems strongly applicable.

Still with the second set, we wonder if there is still potential for automated feedback. In particular, we think the methodology has potential for those courses where term definitions and relationships to other terms is more ambiguous. In this case, our match metric could be used to find similarity over a set of experts and learners. For example, instead of a single criterion map, one could imagine a number of instructors creating maps and similarly students creating mind maps. In addition to these maps, the creator might write a paragraph or two describing why they created the map the way they did. With this database of mind maps, we could run a comparison to a new map and provide that student with a weighted match and show why and how the map they created relates to other learners and experts. This might even be a more meaningful form of feedback than a simpler yes/no response that indicates if you're making a particular mind map relationship.

Looking back at table I the last two columns looked at some simple graph properties for the criterion map. In particular, another question we have is should the words in the criterion map be selected in such a way to manipulate the density or degree of a graph. In simpler language, if the criterion map is simpler (as in less connections between words) is that better for our automatic feedback methodology compared to a a more heavily connected criterion graph. Our results, are unclear at this time, but we believe this is an interesting direction to pursue, where, currently, we hypothesize that a simpler graph will produce better results, but too simple a graph might be detrimental. A follow on question for this hypothesis, is how does an instructor create a criterion map that is appropriate for this methodology, and we leave this as future work.

Overall, in this study we took our early ideas on automatic feedback for student created mind maps to a larger and more diverse set of classes. This study concludes that our approach seems to be useful for courses that use a set of terms that are clearly defined and are used throughout the class. Our future work will focus on refining these techniques by focusing on how to create the criterion map, which in turn determines the subset of terms that will be used in the exercise. The summary data and mind maps are included in a compressed file and can be downloaded at: www.users.miamioh.edu/jamiespa/data_sets/fie_2015_mind_map_data.zip.

References

[1] P. Jamieson, "Using modern graph analysis techniques on mind maps to help quantify learning," in *Frontiers in Education Conference (FIE), 2012*, oct. 2012. [Online]. Available: http://www.users.muohio.edu/jamiespa/html_papers/fie_12.pdf

[2] ——, "More graph comparison techniques on mind maps to provide students with feedback," in *Frontiers in Education Conference (FIE), 2013*, oct. 2013. [Online]. Available: http://www.users.muohio.edu/jamiespa/html_papers/fie_13.pdf

[3] P. Jamieson and J. Eaton, "owards a better graphlet-based mind map metric for automating student feedback," in *to be published at 2015 ASEE Annual Conference and Exposition*, 2015.

[4] T. Buzan and B. Buzan, "The mind map book how to use radiant thinking to maximise your brain's untapped potential," *New York: Plume*, 1993.

[5] T. Anderson, *Study skills and learning strategies*. Center for the Study of Reading, University of Illinois at Urbana-Champaign, 1978. [Online]. Available: http://books.google.com/books?id=JNHUjwEACAAJ

[6] T. Angelo and K. Kross, *Classroom Assessment Techniques - A Handbook for College Teachers*. Jossey-Bass, 2003.

[7] M. J. Somers, K. Passerini, A. Parhankangas, and J. Casal, "Using mind maps to study how business school students and faculty organize and apply general business knowledge," *The International Journal of Management Education*, vol. 12, no. 1, pp. 1–13, 2014.

[8] H. Faste and H. Lin, "The untapped promise of digital mind maps," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2012, pp. 1017–1026.

[9] I. Abi-El-Mona and F. Adb-El-Khalick, "The influence of mind mapping on eighth graders science achievement," *School Science and Mathematics*, vol. 108, no. 7, pp. 298–312, 2008. [Online]. Available: http://dx.doi.org/10.1111/j.1949-8594.2008.tb17843.x

[10] E. Evrekli, D. Inel, and A. Galim, "Development of a scoring system to assess mind maps," *Procedia - Social and Behavioral Sciences*, vol. 2, no. 2, pp. 2330 – 2334, 2010. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S187704281000371X

[11] P. Kedaj, J. Pavlíček, P. Hanzlík *et al.*, "Effective mind maps in e-learning," *Acta Informatica Pragensia*, vol. 3, no. 3, pp. 239–250, 2014.

[12] J. Beel, S. Langer, M. Genzmehr, and B. Gipp, "Utilizing mind-maps for information retrieval and user modelling," in *Proceedings of the 22nd Conference on User Modelling, Adaption, and Personalization (UMAP)*, 2014.

[13] H. E. Herl, H. F. O'Neil, G. K. W. K. Chung, and J. Schacter, "Reliability and validity of a computer-based knowledge mapping system to measure content understanding," *Computers in Human Behavior*, vol. 15, no. 3-4, pp. 315–333, May 1999. [Online]. Available: http://dx.doi.org/10.1016/S0747-5632(99)00026-6

[14] M. Ruiz-Primo, R. Shavelson, and S. Schultz, "On the validity of concept-map-based assessment interpretations: An experimental testing the assumption of hierarchical concept maps in science," University of California, Tech. Rep., 1997. [Online]. Available: http://research.cse.ucla.edu/Reports/TECH455.PDF

[15] I. Nation, *Learning Vocabulary in Another Language*. Cambridge University Press, 2002. [Online]. Available: http://books.google.com/books?id=rg8Aue4jJWkC

[16] A. Coxhead, *Essentials of teaching academic vocabulary*. Houghton Mifflin Company, 2006.

[17] T. Chung and P. Nation, "Technical vocabulary in specialised texts," *Reading in a foreign language*, vol. 15, no. 2, pp. 103–116, 2003.

[18] G. P. Wiggins, J. McTighe, L. J. Kiernan, and F. Frost, *Understanding by design*. Association for Supervision and Curriculum Development Alexandria, VA, 1998.

[19] N. Przulj, D. G. Corneil, and I. Jurisica, "Modeling interactome: scale-free or geometric?" *Bioinformatics*, vol. 20, no. 18, pp. 3508–3515, 2004.