

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/248607319>

Development of a scoring system to assess mind maps

Article in *Procedia - Social and Behavioral Sciences* · December 2010

DOI: 10.1016/j.sbspro.2010.03.331

CITATIONS

77

READS

1,131

3 authors, including:



Didem Inel Ekici
Usak University

51 PUBLICATIONS 613 CITATIONS

[SEE PROFILE](#)



Ali Günay Balım
Dokuz Eylül University

83 PUBLICATIONS 1,160 CITATIONS

[SEE PROFILE](#)

WCES-2010

Development of a scoring system to assess mind maps

Ertuğ Evrekli^{a*}, Didem İnel^b, Ali Günay Balım^c^a*Faculty of Education, Celal Bayar University, Manisa, 45900, Turkey*^b*Faculty of Education, Usak University, Usak, 64200, Turkey*^c*Faculty of Education, Dokuz Eylül University, Izmir, 35150, Turkey*

Received October 20, 2009; revised December 29, 2009; accepted January 11, 2010

Abstract

The present study deals with the use of mind maps as an assessment tools. In the study, the mind maps prepared for the special teaching methods course by 30 pre-service teachers studying in the Department of Science Teacher Training of a university in Turkey in academic year 2008-2009 were assessed by using a scoring system to assess mind maps for pre-service science teachers. To ensure the reliability of the scoring system, the mind maps prepared by the pre-service teachers for the special teaching methods course were assessed by two expert raters and the assessment was repeated one week later. The reliability process for the scoring system was calculated by using inter- and intra-rater reliability values, intra-cluster correlation analysis, and variance analysis.

© 2010 Elsevier Ltd. Open access under [CC BY-NC-ND license](#).

Keywords: Mind maps; science education; assessment; inter-rater and intra-rater agreement.

1. Introduction

Organization of learning environments on the basis of the constructivist approach instead of the behaviorist approach has led to an increase in the number of studies on the techniques and visual materials used for the visual presentation of information. As the constructivist approach suggests that learners structure new information by associating it with their previous knowledge and experiences, it is of great importance to identify the previous knowledge and experiences of students in learning environments. Therefore, studies have been recently conducted on many visual materials in order to reveal students' previous knowledge, to assess what is learned, to attract students' attention to lessons, and to ensure a better retention of the information learned. One of such visual materials is the mind maps. Mind mapping can be described as a visual technique that presents the knowledge, ideas, concepts and the relationships between them in an individual's mental construction on a two-dimensional plane. Balım, Evrekli and Aydın (2007) argue that mind mapping is an effective brain-based visual technique that helps individuals actively use their right brains as well as their left brains by using their associations of the concepts and ideas they have about a central concept or idea and the elements of image, expression, shape, size, and color. Developed by Tony Buzan toward the end of the 1960s, mind maps have been employed in many different areas since their development; yet, they have also become a subject of educational research recently.

* Ertuğ Evrekli. Tel.: +90 236 462 24 88-162; Fax: +90 236 462 16 00

E-mail address: eevrekli@gmail.com

The studies on mind maps are much limited. A review of the studies on mind mapping revealed the following studies aiming to determine the effect of mind maps on achievement (Ammar, 2005; Treviño, 2005; Abi-El-Mona and Abd-El-Khalick, 2008; Akınoğlu and Yaşar, 2007; Bütüner and Gür, 2008), on writing skills (Ling, 2004), on the attitude toward courses (Akınoğlu and Yaşar, 2007), and on recalling (Farrand, Hussain and Hennessy, 2002). Furthermore, Goodnough and Woods (2002) obtained students' and teachers' opinions about mind maps in their study. In their study, Evrekli, Balım and İnel (2009) attempted to identify pre-service teachers' opinions about the use of mind maps in science courses. In their study on medical students, D'Antoni, Zipp and Olson (2009) examined the inter-rater reliability of a rubric to assess mind maps. An examination of the relevant literature demonstrates that although concept maps have been used as an assessment tool in different areas, the studies on the use of mind maps as an assessment tool is scarce. Thus, the present research was carried out.

2. Method

In this study, the mind maps prepared for the special teaching methods course by 30 pre-service teachers studying in the Department of Science Teacher Training of a university in Turkey in academic year 2008-2009 were assessed by using a scoring system to assess mind maps for pre-service science teachers. The scoring systems proposed on concept maps (Novak and Gowin, 2002) and on mind maps were taken into consideration while developing the scoring system (D'Antoni, Zipp and Olson, 2009). The scoring system used to assess mind maps is as follows (Table 1-Figure 1):

Table 1. Scoring system

1 st level concept links (2 point for each if valid)
2 nd level concept links (4 point for each if valid)
3 rd level concept links (6 point for each if valid)
4 th level concept links (8 point for each if valid)
Cross links (10 point for each if valid)
Examples (1point for each if valid)
Relationships (3 point if valid)
Picture, Image and Figure (3 point if valid)
Invalid component (0 point)

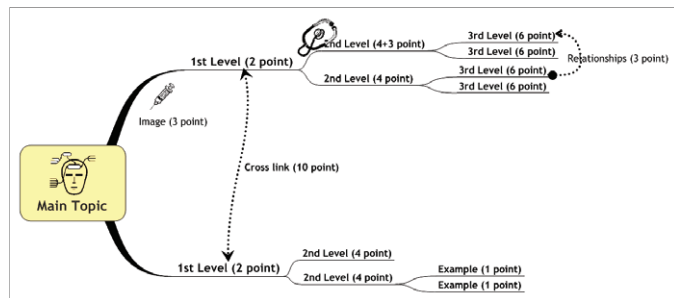


Figure 1. An example about scoring mind maps (in Mind Manager programme)

To ensure the reliability of the scoring system, the mind maps prepared by the pre-service teachers for the special teaching methods course were assessed by two expert raters. Following the expert analyses, to identify the statistical technique suitable for the data set, the Shapiro-Wilk test was performed to determine whether the weighted data had normal distribution for all groups and sub-dimensions. The test results showed that total scores had a normal distribution in four groups ($p > .05$); however, the sub-dimensions did not distribute normally, as revealed by a one-by-one examination of the sub-dimensions ($p < .05$). Therefore, two methods used for continuous data with normal distribution, intra-cluster correlation analysis and Pearson product-moment correlation analysis were performed to determine the agreement between the total scores, while Spearman's rank correlation and Kendall tau-b statistics were employed to determine the agreement between the sub-dimensions. Moreover, variance analysis for repeated measurements was used to ascertain whether there was a significant difference between the intra- and inter-rater assessments on total scores.

3. Findings and Interpretations

This section presents the findings obtained from the study and their interpretations. The study first attempted to calculate the intra-rater reliability for the raters' mind map assessments and thus, assessments were repeated a week later by both raters.

Intra-rater reliability of the scoring system: Agreement between the assessments was determined by using non-parametric statistical techniques for the sub-dimensions and parametric statistical techniques for the total scores. Table 2 shows the findings about the intra-rater reliability for the first and second raters.

Table 2. Intra-rater agreement values for the first and second raters

Sub-dimensions	Rater 1				Rater 2			
	Spearman's ρ		Kendall's τ_b		Spearman's ρ		Kendall's τ_b	
	ρ	P	τ_b	p	ρ	P	τ_b	p
<i>Concept links point</i>	.99	.000<.05	.98	.000<.05	.98	.000<.05	.94	.000<.05
<i>Cross links</i>	.97	.000<.05	.96	.000<.05	.97	.000<.05	.96	.000<.05
<i>Examples</i>	1.00	.000<.05	1.00	.000<.05	.94	.000<.05	.93	.000<.05
<i>Relationships</i>	.99	.000<.05	.98	.000<.05	.98	.000<.05	.96	.000<.05
<i>Image, picture, symbol etc.</i>	.96	.000<.05	.91	.000<.05	.95	.000<.05	.90	.000<.05

The findings obtained from the first rater's mind map analyses suggest that there is a high-level correlation between two measurements for all sub-dimensions and that this correlation is significant ($p=.000<.05$). Furthermore, the analyses also demonstrated that the agreement between the scoring of images, symbols, and other visuals, and examples and the scoring of cross links was relatively lower than the other sub-dimensions. Intra-cluster correlation analysis was performed to compare the repeated analyses on total scores. Şencan (2005) and Shoukri (2004) recommend intra-cluster correlation analysis in determining inter-rater agreement in continuous measurements with equal intervals or proportions. Shrout and Fleiss (1979) proposed three different models for the use of intra-cluster correlation analysis in rater reliability. Given the models, the third intra-cluster analysis model based on the two-way mixed model was used to determine inter-rater agreement (Şencan, 2005). As a result of the intra-cluster analysis performed on the total scores for the first rater's intra-rater reliability, the correlation was calculated as .995 (95% CI=.989-.997), while the intra-cluster correlation for the second rater's intra-rater reliability was .987 (95% CI=.973-.994). Moreover, variance analysis for repeated measurements was used to test whether there was a significant difference between the two scorings for each of the two raters. The results of the analyses are given in Tables 3 and 4.

Table 3. Results of the variance analysis for repeated measurements for the total scores obtained from the two measurements of the first rater

Source of Variance	Sum of Squares	Sd	Mean square	F	p	η^2
Between-subject	189083,9	29	6520,136			
Measurement	13,067	1	13,067	.732	.399	.025
Error	517,933	29	17,860			
Total	189614,9	59				

Total scores obtained from the first rater's assessments were tested using variance analysis for repeated measurements. The analyses revealed no significant difference between the two measurements ($F_{(1-29)}=.732$, $p=.399>.05$). Furthermore, the eta square effect size value calculated also confirms that the independent variable (measurements) did not cause a significant difference upon the dependent variable (score). Moreover, the Pearson product moment for the two measurements was computed to be .995. Table 3 presents the results of the variance analysis for the repeated measurements of the second rater.

Table 4. Results of the variance analysis for repeated measurements for the total scores obtained from the measurements of the second rater

Source of Variance	Sum of Squares	Sd	Mean square	F	p	η^2
Between-subject	186592,923	29	6434,239			
Measurement	5,400	1	5,400	.127	.724	.004
Error	1231,600	29	42,469			
Total	187829,923	59				

According to the results of variance analysis given in Table 3, there was no significant difference between the two repeated measurements of the second rater ($F_{(1-29)}=.127$, $p=.724>.05$). Furthermore, the eta square effect size value calculated demonstrates that the independent variable had a low-level effect on the dependent variable. The Pearson product moment for the two measurements was computed to be .988.

Inter-rater reliability of the scoring system: To calculate the inter-rater reliability of the scoring system, the means of the two raters' measurements were taken for comparison. The analyses on the inter-rater reliability of the scoring system proposed to assess mind maps were carried out by using non-parametric statistical techniques since the scores obtained from the sub-dimensions did not display a normal distribution in each of the two groups. The Shapiro-Wilk's test based on total scores revealed a near-normal distribution of the scores. Therefore, intra-cluster correlation analysis and dependent groups t-test were performed to analyze total scores. Table 5 shows the inter-rater reliability results for the sub-dimensions of the scoring system.

Table 5. Inter-rater agreement for the sub-dimensions of the scoring system

Sub-dimensions	Spearman's ρ		Kendall's τ_b	
	ρ	p	τ_b	P
<i>Concept links point</i>	.94	.000<.05	.86	.000<.05
<i>Cross links</i>	.96	.000<.05	.89	.000<.05
<i>Examples</i>	.77	.000<.05	.74	.000<.05
<i>Relationships</i>	.83	.000<.05	.74	.000<.05
<i>Image, picture, symbol etc.</i>	.91	.000<.05	.81	.000<.05

As a result of inter-rater agreement calculations performed for the scoring of mind maps, it was determined that inter-rater agreement was relatively lower for the scoring of examples, relationships, images, symbols and similar elements in particular, when compared to the other sub-dimensions. Furthermore, in the intra-cluster correlation analyses on total scores, inter-rater agreement was calculated to be .967 (95% CI= .932-.984). Table 6 gives the results of the dependent groups t-test based on the means of the two measurements for each of the two raters.

Table 6. Dependent groups t-test for the mean scores obtained from the two raters

	N	Mean	Standard Deviation	t	p
Rater Mean A	30	124.87	57.10	1.06	.298
Rater Mean B	30	122.03	56.72		

The dependent groups t-test carried out to determine whether there was any significant difference between the raters' mean scores revealed no significant difference between the raters ($t=1.06$; $p=.298>.05$). What is more, the correlation between the two data sets was calculated as .967 according to the Pearson product moment analysis.

4. Discussion, Conclusion and Recommendation

The study proposed a scoring system to assess mind maps and attempted to determine the intra- and inter-rater reliability values for the scoring system. As a result of the analyses, all values regarding the entire scoring system and its sub-dimensions were found to be at a normal level in intra- and inter-rater reliability calculations ($p<.001$). Furthermore, in the calculations, repeated measurements done by each rater were tested by variance analysis for repeated measurements and no significant difference was found between the measurements. As for inter-rater agreement computations, the means were calculated for the scores obtained from the raters' repeated measurements and they were found to agree at a significant level both in the sub-dimensions and total scores. Moreover, dependent groups t-test was performed to test any significant difference between the means, but the test revealed no significant difference. It is believed that this study may provide researchers with an insight into the reliability of the scoring system proposed to assess mind maps. Although the literature abounds with studies on the assessment of concept maps and their use as an assessment tool, studies on mind maps are scarce. A study that is similar to the present one was conducted by D'Antoni, Zipp and Olson (2009). In their study, the researchers proposed a scoring system to assess mind maps and examined the inter-rater reliability of their scoring system. As a result of their study, they obtained a significant agreement between the scorings for calculating colors, images, examples and cross links, while they found no significant agreement for conceptual relationships and hierarchies. Moreover, they determined the agreement value to be .86 as a result of the intra-cluster correlation analyses they performed on total scores. The agreement found by the researchers between the weighted scores calculated on total scores, cross links, examples,

images and colors displays a similarity with the present study. Yet, it could be argued that more studies are needed on the subject. As a result of the study results and literature review, it is believed that:

- The scoring system proposed to assess mind maps can be used for this purpose,
- There is a need for further research about the use of mind maps as assessment tools, and
- The scoring system proposed here could be used in different sample groups and may contribute to obtaining more generalizable results about intra- and inter-rater agreement.

References

- Abi-El-Mona, I. and Adb-El-Khalick. (2008). The Influence of Mind Mapping on Eighth Graders' Science Achievement. *School Science and Mathematics*, 108(7), 298-312.
- Akinoglu, O. and Yaşar, Z. (2007). The Effects of Note Taking in Science Education Through the Mind Mapping Technique on Students' Attitudes, Academic Achievement and Concept Learning. *Journal of Baltic Science Education*, 6(3), 34-43.
- Amma, C. (2005). *Effectiveness of Computer Based Mind Maps in the Learning of Biology at the Higher Secondary Level*. New Delhi: ICDE International Conference (19-23 November).
- Balım, A. G., Evrekli, E. and Aydın, G. (2007). *Fen ve Teknoloji Öğretiminde Zihin Haritalama Tekniği ve Mind Manager Programı Uygulamaları [The Application of Mind Manager Program and Mind Mapping Technique in Science and Technology Education]*. Famagusta, Turkish Republic of Northern Cyprus: VII. International Educational Technologies Conference (3-4-5. May).
- Bütünler, S. Ö. and Gür, H. (2008). Açılar ve Üçgenler Konusunun Anlamlı Öğrenme Araçlarından V Diyagramları ve Zihin Haritaları Kullanılarak Öğretimi [Teaching of Angles and Triangles by using Vee Diagrams and Mind Maps]. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education*, 2(1), 1-18.
- D'Antoni, A. V., Zipp, G. P. and Olson, V. G. (2009). Interrater Reliability of the Mind Map Assessment Rubric in a Cohort of Medical Students. *BMC Medical Education*, 19(9), 1-8.
- Evrekli, E., Balım, A. G. and İnel, D. (2009). Mind Mapping Applications in Special Teaching Methods Courses for Science Teacher Candidates and Teacher Candidates' Opinions Concerning the Applications. *Procedia Social and Behavioral Sciences*, 1, 2274-2279.
- Farrand, P., Hussain, F. and Hennessy, E. (2002). The Efficacy of the Mind Map Study Technique. *Medical Education*, 36, 426-431.
- Goodnough, K. and Woods, R. (2002). *Student and Teacher Perceptions of Mind Mapping: A Middle School Case Study*. New Orleans, LA: The Annual Meeting of the American Educational Research Association (1-5 April).
- Ling, C. W. (2004). *The Effectiveness of Using Mind Mapping Skills in Enhancing Secondary One and Secondary Four Students' Writing in CMI School*. Unpublished Master Thesis, The University of Hong Kong.
- Novak, J. D. and Gowin, D. B. (1984). *Learning How to Learn*. United States of America: Cambridge University Press.
- Shrout, P. E. and Fleiss, J. L. (1979). Intraclass Correlations: Uses in Assessing Rater Reliability. *Psychological Bulletin*, 86(2), 420-428.
- Shoukri, M. M. (2004). *Measures of Interobserver Agreement*. Boca Raton: Chapman and Hall/CRC Press.
- Şencan, H. (2005). *Sosyal ve Davranışsal Ölçümlerde Güvenilirlik ve Geçerlilik [Reliability and Validity in Social and Behavioral Measurement]*. Ankara: Seçkin Yayıncılık.
- Treño, C. (2005). *Mind Mapping and Outlining: Comparing Two Types of Graphic Organizers for Learning Seventh-Grade Life Science*. Unpublished PhD Thesis, Texas Tech University.