Frontiers of Contemporary Education
ISSN 2690-3520 (Print) ISSN 2690-3539 (Online)
Vol. 2, No. 1, 2021
www.scholink.org/ojs/index.php/fce

Original Paper

A Systematic Review of Cognitive Diagnostic Assessment and Modeling through Concept Mapping

Jingshun Zhang[1*], Eunice Jang[2] & Saad Chahine[2]

[1] Jingshun Zhang, Department of Leadership, Technology, and Research, Florida Gulf Coast University, Fort Myers, FL, USA

[2] Eunice Jang, Department of Applied Psychology and Human Development. OISE, University of Toronto, Toronto, ON, Canada

[3] Saad Chahine, Faculty of Education, Queen's University, Kingston, ON, Canada

[*] Jingshun Zhang, E-mail: jzhang@fgcu.edu

Received: December 23, 2020 Accepted: January 4, 2021 Online Published: April 8, 2021

doi:10.22158/fce.v2n1p41 URL: http://dx.doi.org/10.22158/fce.v2n1p41

Abstract

Traditional assessments are typically constructed on logical taxonomies and content specifications but lack explicit cognitive models of the processes and problem-solving strategies that underlie student performance. Cognitive Diagnostic Assessment (CDA) fills this gap by combining cognitive science and psychometrics. CDA is in its infancy, but over 1,000 relevant studies have been conducted in this area during the last 20 years. Facing these complicated studies, many beginners struggle to understand the whole picture of CDA. This paper systematically reviews the literature on CDA and relevant cognitive diagnosis models (CDMs) with the application of a concept mapping technology. Concept mapping is graphical representation of concepts and their relationships. Its use in this study allows researchers and students to gain in-depth knowledge about CDA and CDM and identify areas of future research.

Keywords

literature review, cognitive diagnostic assessment and modeling, concept mapping

1. Introduction

The history of assessment can be traced back to the Chinese civil service in 2200 B.C. (Bowman, 1989), ancient Greeks' individual difference testing (Doyle, 1974), and the European governments' test systems of the 19th century. Testing received a great boost in the 20th century due to rising interest in individual differences (Kline, 2005). In the United States especially, spurred by the *A Nation at Risk*

report in 1983 and the No Child Left Behind Act in 2001, the role of Large-Scale Assessment (LSA) in public education grew tremendously beginning in the mid-1980s, flourished during the 1990s, and unquestionably continued to grow with the implementation of assessment and accountability requirements (DePascale, 2003; Kifer, 2001). Today, assessments of student performance are more prevalent than at any other period in history. National and international assessments have become increasingly frequent as has the use of college admission and placement tests. The development of LSA has been an ongoing process with many aspects still in need of improvement. For example, "the contents of standardized tests and the examination can assess only limited forms of competence, and teachers are quite able to predict which aspects of competence will be assessed" (Wiliam, 2001, p. 165). LSA is an important component of determining achievement in education. It should be carried out in a systematic fashion (Taylor & Tubianosa, 2001) and shows the general public the effectiveness of students, schools, and teachers where there is increased accuracy and objectivity in examining achievement (Crundwell, 2005).

Cognitive(ly) Diagnostic Assessment (CDA) makes explicit the cognitive models of learning beyond the logical taxonomies and discrete content specifications that are prevalent in traditional LSAs. CDA is designed to measure specific knowledge structures and processing skills in students to provide information about their cognitive strengths and weaknesses; it provides formative diagnostic feedback through fine-grained reporting of test takers' skill mastery profiles (Buck & Tatsuoka, 1998; DiBello et al., 1995; Jang, 2009a; Javidanmehr & Anani Sarab, 2017; Nichols, Chipman, & Brennan, 1995). Cognitive Diagnostic Models (CDMs) have been developed "to help construct diagnostic assessments and estimate students' attribute mastery patterns associated with different cognitive skills" (Cui et al., 2006, p. 4). The rapid advance and a growing interest in CDA have resulted in extensive research and theories in the past two decades. Because of its relative infancy and technical complexities, researchers interested in CDA may find it daunting to comprehend the entire picture of this body of research. In this project, we analyzed the relevant studies on CDA and CDMs using concept maps. Concept maps are used to visually represent concepts in a body of knowledge and their relationship to each other. Concept maps show the hierarchy of these concepts and emphasize the links between them. They are valuable pedagogical tools (Cañas et al.; Clark & James, 2004; Novak, 1990).

The purpose of this study was to conduct a systematic review of CDA and its accompanying statistical modelling approaches through concept mapping. This paper specifically focuses on the application of a particular concept mapping technology to the literature on CDA over the last four years. It demonstrates a systematic approach to mapping the multiple areas of interests underlying the research on CDA. Through graphical representation of those areas, the study results shed light on emerging issues that require further development in research on CDA.

2. Literature Review

CDA is designed to measure specific knowledge structures and processing skills that students have acquired and to provide fine-grained diagnostic information about their cognitive strengths and weaknesses (DiBello, Stout, & Roussos, 1995; Jang, 2009; Nichols, Chipman, & Brennan, 1995). CDA can give instructors and policy makers more information about the strategies students use to attack problems, the relationships students perceive among concepts, and the principles students understand in a domain. It provides useful information for educators and policymakers to draw conclusions about students' instructional needs and the effectiveness of instructional programs (Nichols, 1994). By measuring different cognitive processes and the knowledge required to solve test items in a domain of interest, CDA can provide a profile of students' mastery and non-mastery of particular cognitive skills. CDA can have a significant impact on the improvement of instruction and learning by giving students and teachers reliable feedback on student strengths and weaknesses (Roussos, Templin, & Henson, 2007) and "further help design effective interventions for individual students" (Cui et al., 2006, p. 4). CDA provides relevant procedures for application such as cognitive foundations of structured item response models, the Fusion model skill diagnosis system, diagnostic inferences about cognitive skills, and cognitive-diagnostic score reporting.

CDA is "still in its infancy, but its parentage is fairly well established" (Leighton & Gierl, 2007, p. 1) because it can be traced back to the 1950s when prominent researchers, such as Messick, Snow, and Lohman, put these perspectives forward. Over the past three decades, a growing number of researchers have argued that cognitive science and psychometrics should be tightly aligned in the service of instruction (Messick, 1984; Snow & Lohman, 1989). In the 1990s, some influential work was published in books and journals, such as *A Framework for Developing Cognitively Diagnostic Assessments* (Nichols, 1994) and *Cognitively Diagnostic Assessment* (Nichols, Chipman, & Brennan, 1995). Recently, *Cognitive Diagnostic Assessment for Education: Theory and Applications* (Leighton & Gierl, 2007) and *Diagnostic Measurement: Theory, Methods, and Applications* (Rupp, Templin, & Henson, 2010) have become important sources of new ideas leading to innovations in CDA. Much attention has been paid to developing psychometric models that are capable of estimating students' cognitive profiles. CDA has also drawn much attention from practitioners and policy makers who strive for specific information about students' knowledge states and proficiencies that are necessary for making decisions about the allocation of resources as well as planning effective instructional interventions for individual students (Cui et al., 2006; Nichols, 1994; Roussos, Templin, & Henson, 2007). There have also been many studies of relevant theories and models, such as the Fusion model skills diagnosis system, the IRT-based parametric latent class models, the rule space model, the attribute hierarchy method (Cui, 2007; Cui & Leighton, 2008; Cui et al., 2006; Gierl, 2007; Jang, 2008; Roussos et al., 2007) and studies of the broad application of CDA such as skill diagnostics (Jang, 2005, 2009; Jang et al., 2008; Leighton & Gierl, 2007; Zhang, 2013).

During the last decade, evidence in over 600 publications in various forms has shown a rapid growth in the use of CDA (Zhang, Jang, & Chahine, 2011). These studies reported high numbers of broadly classified theoretical debates, psychometric modeling, and applications to a range of testing contexts including both large-scale and classroom assessments when using CDAs. With this rapid growth and CDA's technical complexities, many researchers have expressed a need for a clear framework that can guide test design and validation processes. Despite the large body of published work about CDA, beginning researchers often find it difficult to discern the salient issues that await further research. Therefore, our intent was to examine the status quo of research on CDA and offer a systematic roadmap that could be useful both for developing a much-needed framework of CDA and for beginning researchers.

A systematic literature review is "designed to locate, appraise and synthesize the best available evidence relating to a specific research question in order to provide informative and evidence-based answers" (Boland, Cherry, & Dickson, 2017, p. 2). The systematic review of literature in a narrow area is very important and has "recently become a major area of methodological development" (Gough, Oliver, & Thomas, 2017, p. 9). This idea arose in the 1980s in studies featuring statistical meta-analysis (Glass et al., 1981) and meta-ethnography (Noblit & Hare, 1988). Although the literature review as a form of study has been continuously developed over many years (Boland, Cherry, & Dickson, 2017; Bolin, 2012; Shadish & Lecy, 2015; Sheble, 2017), the systematic literature review is "still a young and rapidly developing field of study and methods of reviewing have not been developed for all areas of science" (Gough, Oliver, & Thomas, 2017, p. 9). There are still many challenges to be overcome (Dickersin, 2015; Oakley et al., 2005) such as conceptual and methodological challenges, resource constraints, the constraints related to the use of reviews, and broader political challenges (Gough, Oliver, & Thomas, 2017). However, many scholars have continued to improve the process used for conducting a systematic literature review in areas such as planning, finding materials to include in the review, describing them, assessing them, analyzing them, synthesizing results, and appraising them (Boland, Cherry, & Dickson, 2017; Lockwood, Munn, & Porritt, 2015; Gough, Oliver, & Thomas, 2012).

This study applied a systematic literature review to the development of CDA to help us to understand its holistic and longitudinal content and relevant theories. Also, this process helped us to enhance a particular research and learning method via concept mapping.

3. Research Methods

3.1 Data Source

We intentionally started to collect relevant references three years ago when we undertook some related studies. After we had accumulated an applicable number of references, we systematically searched them for studies related to CDA. Our research began by locating published sources on CDA research through a university library system. We were able to identify 538 sources in various types of

publications (see Table 1). Among them, 40.5% were journal articles, 17.5% conferences papers, and 16.4% book chapters. Twenty-three books dealt with issues related to CDA, and among them, five focused specifically on CDA or CDM.

Table 1. The Types of Studies in CDA and CDM

Type of studies	Total studies (1957-2010)		Studies in this project (2007-2010)	
	Frequency	Percent	Frequency	Percent
Book	23	4.3	6+2	4.2
Book chapter	88	16.4	27	14.3
Conference	94	17.5	47	24.9
Dissertation	31	5.8	7	3.7
Journal	218	40.5	71	37.6
Report	42	7.8	13	4.2
Workshop	13	2.4	8	4.2
Other	29	5.4	8	4.2
Total	538	100.0	189	100

Initially, we identified sources that spanned the period between 1957 and 2010. Figure 1 shows the distribution of the identified sources. Note that there is a big gap in the years between 1957 and 1973, and that the most significant development of CDA and CDM took place after 2004. Before 2004, the development of CDA and CDM progressed slowly yet steadily.

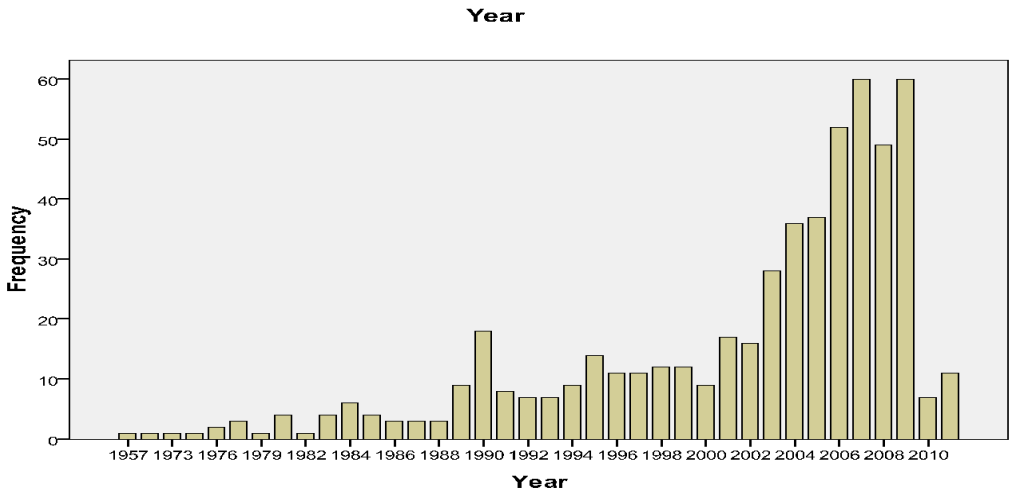

Figure 1. The Distribution of Studies in CDA and CDM

We narrowed our search to focus on sources published between 2007 and 2010, a period that showed a peak in relevant studies. Over the course of these four years, there were 187 sources, which accounted for 34.8% of the total studies from 1957 to 2010. The types of sources are listed in Table 1. These 187 sample sources were reformatted in APA style and numerically coded in a table format. We inputted relevant information about each source, including key words, study contexts, publication type and year, and authors' affiliations.

3.2 Concept Mapping

Alias and Suradi have stated, "Conducting a systematic literature review is an essential research activity in ensuring a good piece of research" (2008, p. 1). A literature review is a multi-stage process that involves scanning information, making notes of reviews, synthesizing and structuring information, writing a critical review of the literature, and building a repertoire of resources (Ali, Mahfouz, & Arisha, 2017; Novak & Cañas, 2006; Rowley & Slack, 2004). Concept mapping can be helpful for conducting literature reviews because it visualizes information using nodes and links arranged in a spatial distribution to reflect the domain information (Alias & Suradi, 2008; Arruarte, Rueda, & Elorriaga, 2008; Carnot, 2006). Rowley and Slack (2004) noted that concept mapping can be useful for identifying key concepts in a collection of documents or a research area, comprehending the structure of a literature review in preparation for writing the review, and understanding theory and concepts as well as the relationships between them. It can serve as a tool for establishing the relationships among areas that are not apparent and are often hierarchical (Novak, 1984). The content, structure, links, and organization of the concept map can be used as the framework for the literature review. There are many kinds of software for concept mapping, both commercial software tools (e.g., 3D Topicscape, Inspiration) as well as free software (e.g., Compendium, Cmap Tools). In this review, a program called Inspiration 9 was used.

4. Result

Based on the relevant literature, the concept mapping in this study took place in three steps: mapping the preliminary structure of CDA and CDM, creating a comprehensive map, and deeply analyzing this map.

In the first stage, we created a preliminary structure map of CDA and CDM based on four main books and four articles (Jang, 2008; Leighton & Gierl, 2007; Nichols, 1994; Nichols, Chipman, & Brennan, 1995; Rupp, Templin, & Henson, 2010; Tatsuoka, 2009). This preliminary structure map included a center (CDA) and four sub-groups, as shown in the following graphic.

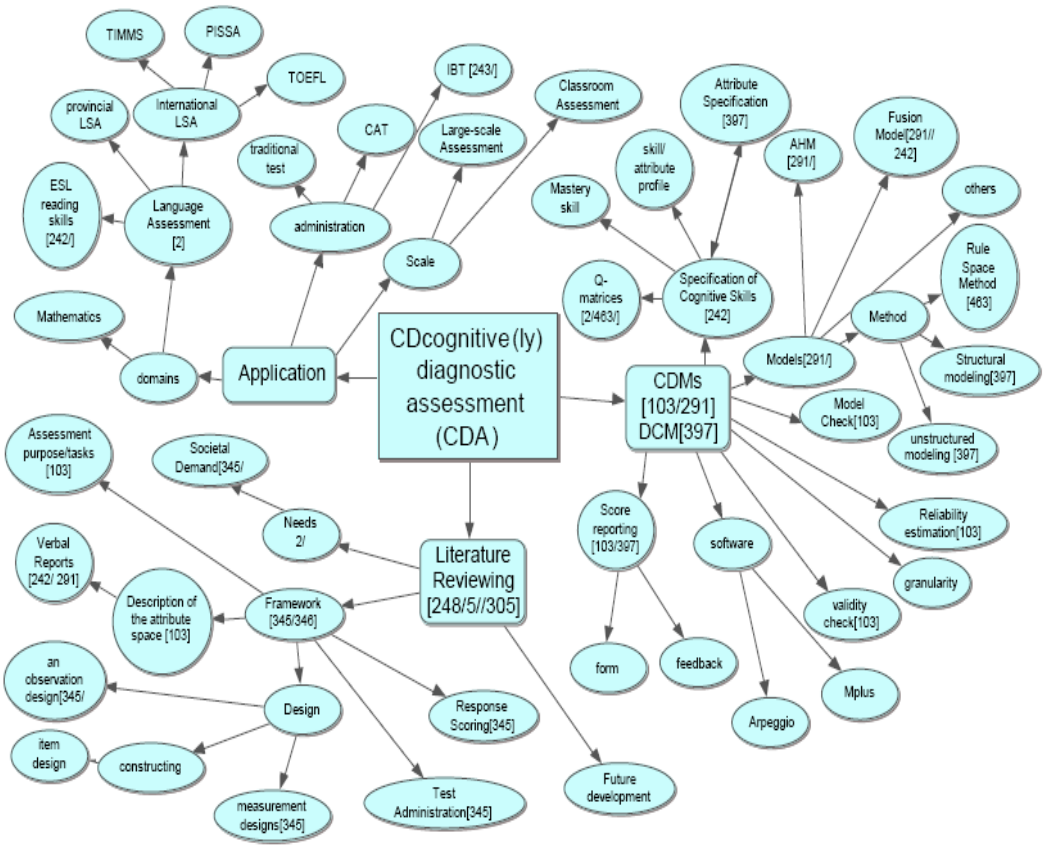

Figure 2. The Preliminary Structure Map

Subsequently, we created a comprehensive CDA and CDM map based on the basic map from step 1. We reviewed all 187 studies and mapped them according to the expanded key words. It was necessary to revisit the original articles to gain some additional information. If an article was involved in a particular area, we would input this study's code into a relevant circle. If there was no relevant circle, a new one would be created based on its relationships with other areas. As a result, the map in Figure 3 successfully represented areas covered by all 187 sources. In our last step, we analyzed each area more substantively and identified distinct issues associated with it.

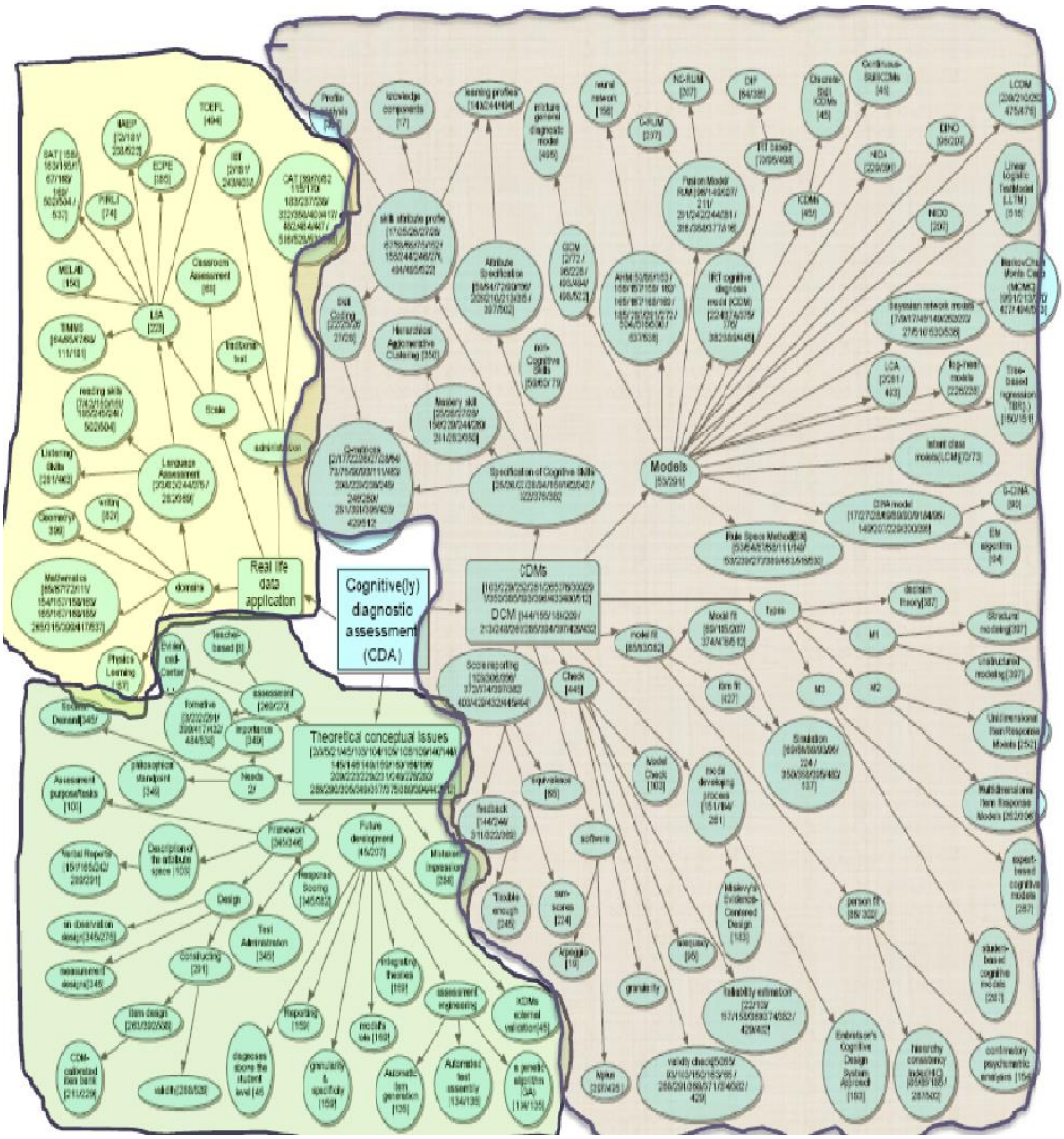

Figure 3. The Comprehensive Map of the Literature Review of CDA and CDM

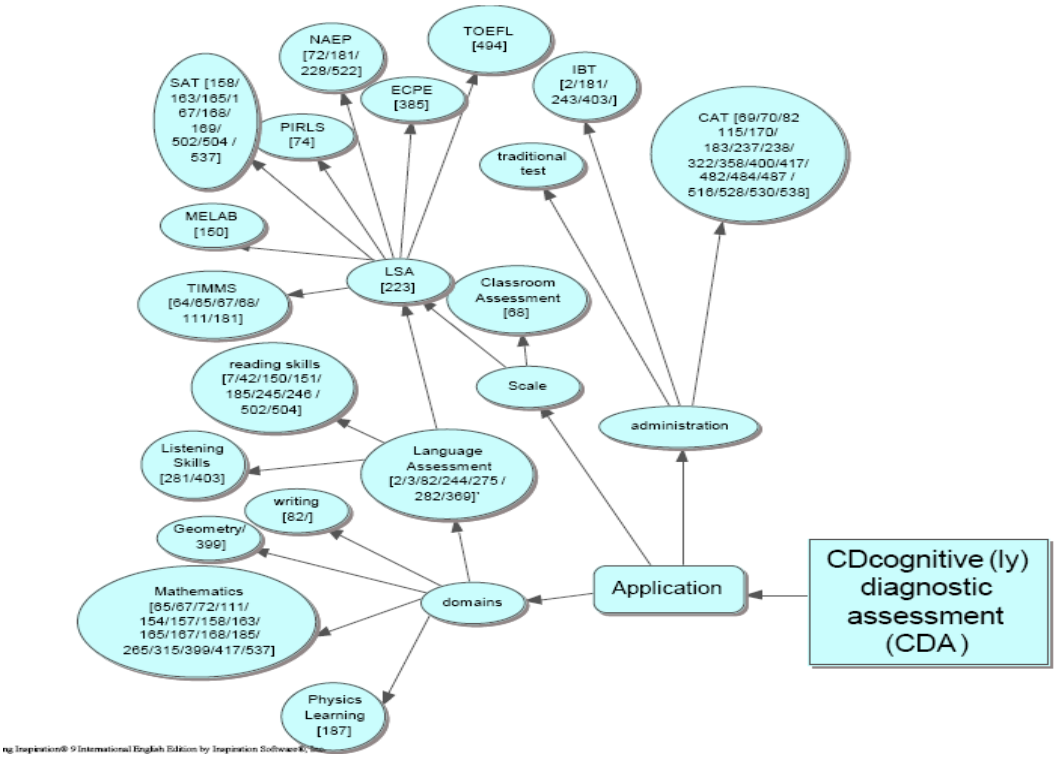

Figure 4. Detail of a Subsection of the Comprehensive Map

After we obtained the comprehensive map in Figure 3, we clarified its hierarchical structures in Microsoft Word format by outputting them from Inspiration 9. For example, we could focus on particular subsections of our comprehensive map, as shown by the dark lines in Figure 3. Figure 4 details the subsection of the comprehensive map that related to the real-life data application of CDA. Table 2 shows its hierarchical structure as it was outputted in Microsoft Word format.

Table 2. Microsoft Word Output of The Hierarchical Structure of Literature in CDA Map

III. Application
- A. Scale
  1. Classroom Assessment [68]
  2. LSA [223]
     - a. TOEFL [494]
     - b. TIMMS [64/65/67/68/ 111/181]
     - c. NAEP [72/181/ 228/522]
     - d. PIRLS [74]
     - e. MELAB [150]
     - f. SAT [158/ 163/165/1 67/168/ 169/ 502/504 /537]
     - g. ECPE [385]

B. Domains
   1. Language Assessment
   [2/3/82/244/275 /282/369]
      a. Reading skills [7/42/150/151/ 185/245/246 /502/504]
      b. Listening Skills [281/403]
   2. Mathematics [65/67/72/111/ 154/157/158/163/ 165/167/168/185/ 265/315/399/417/537]
   3. Writing [82/]
   4. Physics Learning [187]
   5. Geometry/399]

C. Administration
   1. Traditional test
   2. CAT [69/70/82 115/170/ 183/237/238/ 322/358/400/417/ 482/484/487 /516/528/530/538]
   3. IBT [2/181/ 243/403/]

We gained many benefits from the concept mapping process beside the visual maps generated by Inspiration 9. After we created the maps, Inspiration 9 could export additional results in different formats such as PDF, Word, or image files. For example, as mentioned above, Table 2 shows the output of maps in Microsoft Word format. Each file included two parts: at the top of the file, we could retrieve an image of the relevant map; then, the structure of the map became available in text format, which showed the relevant concepts, nodes, and studies in the different categories and levels. This allowed us to recognize the categories and hierarchical structure of the relevant studies; we also could easily track the original studies. In addition, we could establish the areas in which each individual study was involved.

By applying concept mapping, we broadly analyzed the relevant studies on CDA and CDM; we acquired many interesting results. We also considered this process to be a learning experience. By collecting references, analyzing these references, creating maps, and summarizing the maps and relevant outputs, we continued to learn more knowledge and skills that allowed us to adjust our maps continually. Through comparing the many maps made during the process, we found that we also enhanced our research skills and extended our knowledge in more broad areas beside CDA and CDM.

4. Discussion

The project developed over a two-year period. The project included two types of contributions to our CDA development, in theoretical research and in technical support. Theoretically, the results from the application of concept mapping to the literature on CDA offered useful insight into what has and has not been researched. Applying the concept mapping technology to a literature review of CDA helped us to

visually compartmentalize the large body of literature. We categorized the literature into three main classifications: a) real-life data application, b) theoretical and conceptual issues, and c) psychometric modeling (see Figure 2). It appears that much attention has been given to the development of psychometric models, and more research is necessary to examine the validity of CDA in specific learning contexts. The field will benefit from a systematic and comprehensive framework for developing and validating CDA. Two books (Leighton & Gierl, 2007; Rupp, Templin, & Henson, 2010) have done excellent work in this direction.

In the technical and academic support for CDA and CDM, the approach used in this project might prove to be a very useful practice (Zhang, 2016). Concept mapping, as we used it, can be a deductive method for building a conceptual model. As Dr. Jackie Leighton said in the keynote address of the Division D award of the American Educational Research Association (AERA) in 2010, the larger measurement community must develop a conceptual-visual model of CDA to further establish it as an important field; such a model can serve as a method of communication and can open CDA up to the larger research community. This will allow more researchers to understand the literature well and to draw on several sections of it when writing papers. In connection with the above discussions, what we advocate is a deductive approach—instead of thinking up a conceptual model based on our experience or our ideas, we may be able to concept map the literature in order to develop a conceptual model of a theory.

Based on our results, we would suggest the creation of an academic online platform build up a virtual on-line academic community for relevant technical and academic support of CDA and CDM. This online platform might serve some of the following functions:

1) Concept maps of CDA and CDM would create visual images of the relevant studies and theories that researchers and learners could refer to.
2) Reference pools would include links to all part and points in these maps. Users interested in some parts or points could click on them to track the relevant references.
3) Key elements such as reference lists in APA style, key words, and abstracts for each reference could be included in the reference pools. We could post the relevant articles on the web system after receiving copyright permission for their use in this project. The academic online platform, which could be named The Systematic Review of CDA and CDM, could be an open system. We could set up a function that would allow for group discussion. It would be open to all researchers and learners. Members could discuss any question and share ideas after registering to obtain a user account.

Using this system, we hope to build an online and virtual learning community that would be helpful for researchers, learners, and CDA's development in the future (Kernan, Basch, & Cadorett, 2018).

In summary, we conducted the first systematic review of the literature relevant to CDA and CDM using a concept mapping method. This study contributes to the field by applying this new method of literature review to CDA and CDM studies. This study showed this method is useful for building an entire picture of a complex body of research and opening a window to future studies in theoretical and technical support. Posting these maps on-line could provide beginners, researchers, and others with information on

relevant studies and clarify the framework of CDA and CDM.

References

Ali, A., Mahfouz, A., & Arisha, A. (2017). Analysing supply chain resilience: Integrating the constructs in a concept mapping framework via a systematic literature review. *Supply Chain Management: An International Journal*, *22*(1), 16-39. https://doi.org/10.1108/SCM-06-2016-0197

Alias, M., & Suradi, Z. (2008). Concept mapping: A tool for creating a literature review. In A. J. Cañas, P. Reiska, M. Ahlberg, & J. D. Novak (Eds.), *Concept Mapping: Connecting Educators Procedure of the Third Int. Conference*. Tallinn, Estonia & Helsinki, Finland.

Arruarte, A., Rueda, U., & Elorriaga, J. A. (2008). Organizing the Learning Resources Related to the Subject Introduction to Artificial Intelligence through Concept Maps. In *Paper presented in 38th ASEE/IEEE Frontiers in Education Conference, October 22-25, 2008*. Saratoga Springs, NY. https://doi.org/10.1109/FIE.2008.4720484

Boland, A., Cherry, G., & Dickson, R. (2017). *Doing a systematic review: A student's guide* (2nd ed.). SAGE.

Bolin, I. (2012). Formalising syntheses of medical knowledge: The rise of meta-analysis and systematic review. *Perspectives on Science*, *20*(3), 273-309. https://doi.org/10.1162/POSC_a_00075

Buck, G., & Tatsuoka, K. (1998). Application of the rule-space procedure to language testing: Examining attributes of a free response listening test. *Language Testing*, *15*(2), 119-157. https://doi.org/10.1177/026553229801500201

Cañas, A. J., Coffey, J. W., Carnot, M. J., Feltovich, P., Hoffman, R. R., Feltovich, J., & Novak, J. D. (2003). A Summary of Literature Pertaining to the Use of Concept Mapping Techniques and Technologies. Prepare for *The Chief of Naval Education and Training Pensacola FL 32500*.

Carnot, M. J. (2006). Using concept maps to organize information for large scale literature reviews and technical reports: two case studies. In A. J. Cañas, & J. D. Novak (Eds.), *Concept Maps: Theory, Methodology, Technology Proc. of the Second Int. Conference on Concept Mapping*. San José, Costa Rica.

Clark, I., & James, P. (2004). Using concept maps to plan an introductory structural geology course. *Journal of Geoscience Education*, *52*(3), 224-230. https://doi.org/10.5408/1089-9995-52.3.224

Crundwell, R. M. (2005). Alternative strategies for large scale student assessment in Canada: Is value-added assessment one possible answer. *Canadian Journal of Educational Administration and Policy*, *41*, 1-21.

Cui, Y. (2007). *The hierarchy consistency index: Development and analysis*. Unpublished doctoral dissertation at University of Alberta, Edmonton, Alberta, Canada.

Cui, Y., & Leighton, J. P. (2009). The hierarchy consistency index: Evaluating person fit for cognitive diagnostic assessment. *Journal of Educational Measurement*, *46*, 429-449. https://doi.org/10.1111/j.1745-3984.2009.00091.x

Cui, Y., Leighton, J. P., Gierl, M. J., & Hunka, S. (2006, April). *A person-fit statistic for the attribute hierarchy method: The hierarchy consistency index*. Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco, CA.

DePascale, C. A. (2003). The ideal role of large-scale testing in a comprehensive assessment system. *Journal of Applied Testing Technology*, 1-10. Retrieved March 14, 2010, from http://www.testpublishers.org/Documents/Large_Scale_Assessment_v3.0.pdf

DiBello, L. V., Stout, W. F., & Roussos, L. A. (1995). Unified cognitive/ psychometric diagnostic assessment likelihood-based classification techniques. In P. D. Nichols, S. F. Chipman, & R. L. Brennan (Eds.), *Cognitively Diagnostic Assessment* (pp. 361-389). Hillsdale, NJ: Lawrence Erlbaum Associates.

Dickersin, K. (2015). Innovation and cross-fertilization in systematic reviews and meta-analysis: The influence of women investigators. *Research Synthesis Methods*, *6*(3), 277-283. https://doi.org/10.1002/jrsm.1147

Doyle, K. O. Jr. (1974). Theory and practice of ability testing in Ancient Greece. *Journal of the History of the Behavioral Science*, *10*, 202-212. https://doi.org/10.1002/1520-6696(197404)10:2<202::AID-JHBS2300100208>3.0.CO;2-Q

Glass, McGaw, B., & Smith, M. L. (1981). *Meta-Analysis in social research*. Rewbury Park, CA: SAGE.

Gough, D., Oliver, S., & Thomas, J. (2017). *An introduction to systematic reviews* (2nd ed.). SAGE.

Gough, D., Oliver, S., &Thomas, J. (2012). Clarifying differences between designs and method. *Systematic Review*, *1*(28), 1-9. https://doi.org/10.1186/2046-4053-1-28

Jang, E. E. (2008). A framework for cognitive diagnostic assessment. In C. A. Chapelle, Y. R. Chung, & J. Xu (Eds.), *Towards adaptive CALL: Natural language processing for diagnostic language assessment* (pp. 117-131).

Jang, E. E. (2009). Cognitive diagnostic assessment of L2 reading comprehension ability: Validity arguments for applying Fusion Model to Language assessment. *Language Testing*, *26*(1), 31-73. https://doi.org/10.1177/0265532208097336

Javidanmehr, Z., & Anani Sarab, M. R. (2017). Cognitive diagnostic assessment: Issues and considerations. *International Journal of Language Testing*, *7*(2), 73-98.

Kernan, W. D., Basch, C. H., & Cadorett, V. (2018). Using Mind Mapping to Identify Research Topics: A Lesson for Teaching Research Methods. *Pedagogy in Health Promotion*, *4*(2), 101-107. https://doi.org/10.1177/2373379917719729

Kifer, E. (2001). *Large-scale assessment: Dimensions, dilemmas, and policy*. Thousand Oaks, CA: Sage Publications.

Kline, T. J. B. (2005). *Psychological testing: A practical approach to design and evaluation*. Thousand Oaks, CA: Sage Publications. https://doi.org/10.4135/9781483385693

Leighton, J. P., & Gierl, M. J. (Eds.). (2007). *Cognitive Diagnostic Assessment for Education: Theory and Applications*. New York, NY, US: Cambridge University Press. https://doi.org/10.1017/CBO9780511611186

Lockwood, C., Munn, Z., & Porritt, K. (2015). Qualitative research synthesis: methodological guidance for systematic reviewers utilizing meta-aggregation. *International Journal of Evidence-Based Healthcare*, *13*(3), 179-187. https://doi.org/10.1097/XEB.0000000000000062

Messick, S. (1984). The psychology of educational measurement. *Journal of Educational Measurement*, *21*, 215-237. https://doi.org/10.1111/j.1745-3984.1984.tb01030.x

Nichols, P. D. (1994). A framework for developing cognitively diagnostic assessments. *Review of Educational Research*, *64*(4), 575-603. https://doi.org/10.3102/00346543064004575

Nichols, P. D., Chipman, S. F., & Brennan, R. L. (Eds). (1995). *Cognitively diagnostic assessment*. Mahwah NJ: Lawrence Erlbaum Associates.

Noblit, G., & Hare, R. D. (1988). *Meta-ethnography: Synthesizing qualitative studies*. Newbury Park, NY: Sage Publications. https://doi.org/10.4135/9781412985000

Novak, J. D., & Cañas, A. J. (2006). The theory underlying concept maps and how to construct them. In *Technical Report IHMC Cmap Tools 2006-01*. Retrieved 21/6/07, Florida Institute for Human and Machine Cognition, from http://cmap.ihmcus/Publications/ResearchPapers/TheoryUnderlyingConcept Maps.pdf

Novak, J. D., & Gowin, D. B. (1984). *Learning How to Learn*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9781139173469

Roussos, L. A., Templin, J. L., & Henson, R. A. (2007). Skills diagnosis using IRT-based latent class models. *Journal of Educational Measurement*, *44*(4), 293-311. https://doi.org/10.1111/j.1745-3984.2007.00040.x

Rowley, J., & Slack, F. (2004). Conducting a literature review. *Management Research News*, *27*(4), 31-39. https://doi.org/10.1108/01409170410784185

Rupp, A. A., Templin, J., & Henson, R. J. (2010). *Diagnostic measurement: Theory, Methods, and Applications*. New York: Guilford Press.

Shadish, W. R., & Lecy, J. D. (2015). The meta-analytic big bang. *Research Synthesis Methods*, *6*(10), 3246-3264. https://doi.org/10.1002/jrsm.1132

Sheble, L. (2017). Macro-level diffusion of a methodological knowledge innovation: Research synthesis methods, 1972-2011. *Journal of the Association for Information Science and Technology*, *68*(12), 2693-2708. https://doi.org/10.1002/asi.23864

Snow, R. E., & Lohman, D. F. (1989). Implications of cognitive psychology for educational measurement. In R. L. Linn (Ed.), *Educational measurement* (pp. 263-331). New York: American Council on Education, Macmillian.

Tatsuoka, K. K. (2009). *Cognitive Assessment: An Introduction to the Rule Space Method*. New York: Routledge. https://doi.org/10.4324/9780203883372

Taylor, A. R., & Tubianosa, T. (2001). *Student assessment in Canada: Improving the learning environment through effective evaluation*. Kelowna, BC: Society for the Advancement of Excellence in Education.

Wiliam, D. (2001). An overview of the relationship between curriculum and assessment. In D. Scott (Ed.), *Curriculum and Assessment* (pp. 165-181). Westport, CT: Ablex Publishing.

Zhang, J. (2013). *Relationships between missing responses and skill mastery profiles of cognitive diagnostic assessment*. Diss. University of Toronto.

Zhang, J. (2014). *Missing responses, characteristics, and skill mastery profiles of CDA*. Lambert Academic Publishing.

Zhang, J., Jang, E. E., & Chahine, S. (2011, April). *A systematic review of cognitive diagnostic assessment and modeling through concept mapping*. Paper presented at the annual meeting of the American Educational Research Association in 2011, New Orleans, LA.