Evaluation and Program Planning 60 (2017) 166–175

Contents lists available at ScienceDirect

Evaluation and Program Planning

journal homepage: www.elsevier.com/locate/evalprogplan

Introduction to a special issue on concept mapping

William M. Trochim, Professor Policy Analysis & Management[a,*], Daniel McLinden[b]

[a] Cornell University, 435 Kennedy Hall, Ithaca, NY 14853, United States
[b] Cincinnati Children's Hospital Medical Center, United States

ARTICLE INFO

Article history:
Received 22 July 2016
Accepted 7 October 2016
Available online 11 October 2016

Keywords:
Concept mapping
Structured conceptualization
Brainstorming
Sorting
Multidimensional scaling
Cluster analysis
Pattern matching

ABSTRACT

Concept mapping was developed in the 1980s as a unique integration of qualitative (group process, brainstorming, unstructured sorting, interpretation) and quantitative (multidimensional scaling, hierarchical cluster analysis) methods designed to enable a group of people to articulate and depict graphically a coherent conceptual framework or model of any topic or issue of interest. This introduction provides the basic definition and description of the methodology for the newcomer and describes the steps typically followed in its most standard canonical form (preparation, generation, structuring, representation, interpretation and utilization). It also introduces this special issue which reviews the history of the methodology, describes its use in a variety of contexts, shows the latest ways it can be integrated with other methodologies, considers methodological advances and developments, and sketches a vision of the future of the method's evolution.

1. Background on this special issue

This special issue of *Evaluation and Program Planning* is about the reaching of a milestone, of an anniversary worth noting. Work on it commenced when we realized that it was nearly twenty-five years since the publication of a special issue of this same journal (Trochim, 1989c) that is now recognized as the "coming out" of the methodology referred to as "concept mapping" that has evolved steadily over the past decades and has shaped research over that duration in unique and enduring ways. While there were several publications prior to this special issue that touched upon (Trochim, 1985) and introduced (Trochim & Linton, 1986) the concept mapping approach, it was the publication of the 1989 volume that provided the critical push of the methodology into a wider awareness. The first mention of the method (Trochim, 1985) was tangential and the description of it incomplete. The earliest introduction (Trochim & Linton, 1986) went largely unnoticed until after the 1989 volume called attention to it. To date, that earlier paper has had only 49 citations, most either from the special issue commemorated here or from subsequent publications that invariably also cited the special issue. In short, prior to the 1989 special issue, the method was essentially an "insider's" game, known to a very small community of evaluators, faculty and graduate students, who were associated with the graduate program in evaluation in the College of Human Ecology at Cornell University. Despite these historical precedents, it was the 1989 special issue, and especially the lead article in that issue (Trochim, 1989a), that became the standard citation for the concept mapping methodology in the literature. But anniversaries are ephemeral things. Our original intent was to commemorate the 1989 special issue by publishing this follow-up at the 25-year mark in 2014. Alas, the production of such special issues being what they are, that anniversary has already passed, and it may in the end wind up being more reasonable to consider this volume a 30th anniversary of the 1986 origin article.

The purposes of this introductory article are to provide a brief overview of the concept mapping method and how it is accomplished for those who may not be familiar with it, and to explain how the set of articles in this special issue have been organized.

2. What is concept mapping?

2.1. Definition

"Concept mapping is a structured process, focused on a topic or construct of interest, involving input from multiple participants, that produces an interpretable pictorial view of their ideas and concepts and how these are interrelated." (Trochim, 1989b). From the very beginning, there was a strong emphasis on CM being a

* Corresponding author.
E-mail address: wmt1@cornell.edu (W.M. Trochim).

http://dx.doi.org/10.1016/j.evalprogplan.2016.10.006
0149-7189/

"structured" process. This was deliberate and meant as a contrast with approaches to mapping that were unstructured, that produced maps through processes that occurred only in the mind of the producer. CM was to be more transparent than those approaches, and structuring the process was a way to accomplish that end.

Note that this definition required multiple participants. This type of CM is inherently a group process. Some (including others in this volume) have elected to signal that and distinguish this method from many other approaches that use the label "concept mapping" by calling the approach Group Concept Mapping (GCM). But that isn't entirely satisfying either. It leaves open the possibility that it could be construed as trying to represent something like a "group mind" (McDougal, 1920) or it suggests confusion with the negative notion of "group think" (Janis, 1982). Perhaps better would be terms like "participatory" concept mapping or "collaborative" concept mapping. But history may decide the matter. There is a need for a way to distinguish this method from other mapping approaches. The GCM label already has a toehold and, despite problems, may prevail. In this Special Issue we will generally encourage either the use of the traditional abbreviation CM to represent the specific type of concept mapping described here or the increasingly common GCM to serve as an equivalent.

3. Steps in the process

In its original incarnation (Trochim & Linton, 1986) CM was described as involving three steps: (1) the *generation* of the conceptual domain; (2) the *structuring* of the conceptual domain; and (3) the *representation* of the conceptual domain. By the time of the special issue of *Evaluation and Program Planning* (Trochim, 1989 b) these had been further divided into the now-standard (Kane & Trochim, 2006) six-step model: Preparation; Generation; Structuring; Representation; Interpretation; and Utilization. The original presentation omitted the additional three steps because that explication had the primary aim of casting CM in a generalized framework of all structured conceptualization methods. The three steps of preparation, interpretation and utilization represent the "wrap-around" processes within which the heart of the method resides. They constitute the essential facilitation steps to implement the method in a real-world context.

In this introduction to the CM process steps, we'll present only the bare essentials needed to provide an overview of the method to those readers who may be unfamiliar with it. More detailed descriptions are presented in the major book-length treatment of the method (Kane & Trochim, 2006) and in the methods sections of virtually any of the publications on CM. Virtually every step and its many sub-steps encompass choices and there are myriad ways one could address them that are described throughout the literature. This presentation avoids discussing those nuances and shows a very typical way that concept mapping might be instantiated. The hypothetical context we'll use for this presentation is the development of or planning for an imaginary program (the XYZ Program) that might be a treatment or intervention that will eventually be evaluated. We'll assume that there is a trained concept mapping facilitator who helps guide the entire project and that there are one or more client initiators who hire that facilitator as a consultant to help them accomplish it.

3.1. Preparation

One of the most challenging steps in the CM process is simply getting it started. The facilitator has to have an initial person or persons with whom they initiate and plan the project and be able to help them work through the many variations in implementation that are possible. There are several things that need to be addressed as part of the preparation of a CM project. First, one needs to determine the focus of the project. This is usually operationalized as the development of a *focus statement* or *focus prompt* (these are essentially alternate forms of the same underlying focus). Every concept map has a single focus, although to term this a "focus" may incorrectly suggest that it is narrow in scope. The focus statement is worded as an imperative command that describes what the content of the map will attempt to encompass and can be used to guide participants in the generation of statements. For instance, a hypothetical focus statement might read "Generate statements (short phrases or sentences) that describe the specific elements or components of the XYZ program." The focus prompt re-words this focus statement so that it can be the beginning of a complete-the-sentence instruction for brainstorming of statements: "One specific element or component that should be included as part of the XYZ Program is . . . " The focus statement describes the actual focus of the project; the focus prompt is more useful in guiding the generation and editing of the statements.

In addition to the substantive focus of the mapping project one also typically needs to develop the focus for one or more ratings of the statements that will be used in the analysis of maps (especially in pattern matches and go-zone plots). The most typical rating focus is the instruction for a rating of the relative importance of the statements and might look something like:

Please rate each statement for its relative importance where:

1 = relatively unimportant (compared with the rest of the statements)

2 = somewhat important

3 = moderately important

4 = very important

5 = extremely important (compared with the rest of the statements)

It is technically not necessary to have any ratings in order to accomplish concept mapping. The minimally sufficient data needed for a map is the sorting data. But the relative ease and low burden of subjective ratings, and their value in enhancing the richness of the results, typically leads to the inclusion of one or more. In addition to importance, one could rate relative feasibility, cost, influence, priority, and so on. And, it is not necessary that all participants do all ratings. Sometimes a subgroup of participants are in a better position to judge some aspects of the statements than others, so only that subgroup needs to rate those characteristics. In fact, the term "rating" is probably an unnecessarily restrictive one. Virtually any characteristic that can be measured for each statement in the set could be included whether it is a rating or not. For instance, one could calculate the actual costs of accomplishing each statement or the estimates of the pre-post gain on assessments of each statement rather than a formal subjective "rating". However these measurements are accomplished, during the preparation step of the process one needs to determine the procedures that will be followed.

Second, during the preparation of a CM project one needs to identify who the participants in the project will be. There a wide variety of ways one could involve participants. In the extreme case, CM could be accomplished by a very ambitious individual who generates all of the conceptual content, sorts the statements enough times to create a map and rates them with enough specificity and precision to enable reasonable pattern match and go-zone plots. We don't know of any published examples of this one-participant version, but it is theoretically possible. At the other extreme, it is possible to have many different groups and subgroups participate at different points in the process. The people who plan the mapping process may or may not participate. The people who generate ideas (e.g., through brainstorming) may not be the same ones who sort or rate the ideas or the ones who do the interpretation of the map. The flexibility with which people can be involved in the process is a potential strength of the method that requires considerable management by the facilitator and initiator (s) of the process.

As in any applied social research where people are involved, the selection of participants can be construed as a sampling problem. In the typical archetypal concept mapping project we don't sample randomly and don't sample for representativeness of some population (although one could). The purpose of most CM projects is different from sample survey designs where inferences to a population of interest are usually more prominent. In CM, we are typically more interested in sampling from a potential conceptual universe than in sampling people proportionally from a population. For any given topic, the intent in CM (especially in the generation phase) is to represent the topic as accurately and in as much detail as possible. We don't want to miss or overlook any important topics in the conceptual domain. In most CM projects a type of opportunistic sampling for heterogeneity (Trochim, Donnelly, & Arora, 2016) is used. Essentially, we typically want to have a diverse group of participants to help assure that all major perspectives on the topic will be represented, even if they won't be represented in proportion to what exists in the population of participants. In other words, there is an emphasis in CM in assuring that minority perspectives (with respect to the focus) in any context will be included. In most projects an eclectic and opportunistic approach will be taken to identifying the participants, including using members of committees or organizations, email lists, informal networks, word-of-mouth approaches, snowball techniques, and so on.

The other aspect of preparation with respect to participants is to determine what demographic or descriptive data one will collect from them. In CM, demographic variables are used to describe the groups that participate and as variables that enable subgroup analyses of mapping results. Demographic variables range from general descriptors such as gender, race, or income to more context-specific factors like department or degrees obtained.

Third, during the preparation step one needs to plan the logistics for the project, especially the overall project schedule. Although CM has been accomplished in as little as a single day, it typically takes place over weeks or months and involves multiple discrete interactions with participants. For instance, in a contemporary web-based implementation, there may be one or more sessions to introduce the idea to participants, a several week period for web-based generation or brainstorming of statements, several days for synthesis and editing of the brainstormed statements, several weeks for web-based collection of the demographics, sorting and rating data, a period of time for conducting the analysis, one or more scheduled interpretation sessions, and whatever time is required to accomplish utilization of the map results. The logistics planning includes setting this schedule and preparing all of the materials for managing the project including determining how and when participants will be contacted and the drafting of the content for such contacts.

3.2. Generation

The CM method doesn't care how one generates the statements for a mapping project as long as they represent adequately and as completely as possible all of the key facets of the conceptual domain. The most typical way to generate statements is through some form of brainstorming (Coxon, 1999; Osborn, 1948). The primary constraint is that the statements remain true to the focus statement. Depending on the context, one might take extra steps to assure that minority views that might get suppressed in an open live brainstorming are included, such as conducting anonymous brainstorming. But there are many alternatives to brainstorming that could also yield acceptable concept mapping statement sets, including abstraction of text documents through content analysis (Krippendorff, 2004), editing of focus group transcripts (Kreuger, 1988; Krueger & Casey, 2000; Morgan, 1988; Stewart & Shamdasani, 1990), abstraction of open-ended qualitative responses from surveys (Jackson & Trochim, 2002), extraction of key ideas from literature reviews, or even combinations of these.

The intent in generation is to create as rich and varied a set of statements as can be reasonably processed in subsequent steps. Given the data collection demands, especially for the sorting task, this usually means that the final statement set will be between about 80–100 statements. Typically before finalizing the statement set that constitutes the operationalized conceptual domain for the project it is necessary to do an editing and synthesis of the originally edited statements. Web-based brainstorming in particular tends to yield larger statement sets with poorly worded statements and many redundant ideas. These need to be edited for grammar, spelling, and elimination of acronyms and technical language that would not be generally understandable to participants in subsequent steps. It is important to accomplish this synthesis while simultaneously preserving as much of the content and wording in the original voice of the participants as possible. Usually some type of structured editing process or content analysis is used to synthesize the final statement set. One useful rule is that one should keep an audit trail from the original generated ideas to the final edited set so that an outside observer would be able to make an independent judgment of the validity of the decisions that get made in this synthesizing process.

3.3. Structuring

The structuring step in concept mapping is what would typically be referred to in research as the data collection. In this stage one collects the demographic variables, sorting data, and any ratings. For demographics, each participant is asked to provide general non-identifying information that will make it possible to describe the participant groups and classify them into subgroups for more detailed analysis. For the sorting (Coxon, 1999; Rosenberg & Kim, 1975; Weller and Romney, 1988), each participant groups the statements into groups "in a way that makes sense to you." The only restrictions in this sorting task are that there cannot be: (a) N groups (every group having one item each); (b) one group consisting of all items; or (c) a 'miscellaneous' group (any item thought to be unique is to be put in its own separate pile). If software is used it needs to allow the participant to create, delete and name new groups and to move statements from one group to another. Weller and Romney (1988) point out why unstructured sorting (in their terms, the pile sort method) is appropriate for CM data:

> The outstanding strength of the pile sort task is the fact that it can accommodate a large number of items. We know of no other data collection method that will allow the collection of judged similarity data among over 100 items. This makes it the method of choice when large numbers are necessary. Other methods that might be used to collect similarity data, such as triads and paired comparison ratings, become impractical with a large number of items (p. 25).

For the rating task, each participant typically rates each statement on a 5-point Likert-type response scale for each desired rating (e.g., importance, feasibility).

Increasingly data are collected over the web using a standard web browser and proprietary software system, most typically these days the Concept System® Global MAX© (Concept Systems Incorporated, 2005) which is designed to make the collection of concept mapping data over the web easy to accomplish.

3.4. Representation

In CM, the representation step is equivalent to what is meant by data analysis in most social research. The term “representation” was used because it is through the analysis that we represent the data in the form of the results (maps, pattern match and go-zones plots). The analysis begins with construction from the sort information of an NxN binary, symmetric matrix of similarities, Xij. For a single participant and for any two statements i and j, a 1 is placed in Xij if the two items were placed in the same pile by the participant, otherwise a 0 is entered (Weller and Romney, 1988). The total NxN similarity matrix, Tij is obtained by summing across the individual participants’ Xij matrices. Therefore, any cell in this total matrix could take integer values between 0 and the number of people who sorted the statements. The specific values in the total matrix indicate the number of people who placed the i,j pair in the same pile.

The total similarity matrix Tij is analyzed using nonmetric multidimensional scaling (MDS) analysis with a two-dimensional solution. The solution is limited to two dimensions because, as Kruskal and Wish (1978) point out:

> Since it is generally easier to work with two-dimensional configurations than with those involving more dimensions, ease of use considerations are also important for decisions about dimensionality. For example, when an MDS configuration is desired primarily as the foundation on which to display clustering results, then a two-dimensional configuration is far more useful than one involving three or more dimensions (p. 58).

The analysis yields a two-dimensional (x,y) configuration of the set of statements based on the criterion that statements piled together most often are located more proximately in two-dimensional space while those piled together less frequently are further apart.

The x,y configuration is the input for the hierarchical cluster analysis utilizing Ward's algorithm (Anderberg, 1973; Everitt, 1980) as the basis for defining a cluster. Using the MDS configuration as input to the cluster analysis in effect forces the cluster analysis to partition the MDS configuration into non-overlapping clusters in two-dimensional space. There is no simple mathematical criterion by which a final number of clusters can be selected. The procedure that is typically followed is to examine an initial cluster solution that was the maximum desirable for interpretation in this context. Then, successively lower cluster solutions are examined, with a judgment made at each level about whether the merger seems substantively reasonable. The pattern of judgments of the suitability of different cluster solutions is examined and the final number of clusters selected to preserve the most detail and still yield substantively interpretable clusters of statements.

The MDS configuration of the statement points is graphed in two dimensions. This “point map” displays the location of all the brainstormed statements with statements closer to each other generally expected to be more similar in meaning. To illustrate what an initial point map looks like Fig. 1 shows a hypothetical map of a set of 89 statements, with each statement identified by its arbitrary identification number. Throughout this article all example graphs are based on a constructed hypothetical example, not on real CM data.

Several things are noteworthy about any point map. First, no axes are shown even though MDS is used to construct a two-dimensional (x,y axis) configuration because CM focuses on interpreting the relational structure not on an axial interpretation. That is, the primary characteristic of importance for CM is the relationships among the points as indicated in their distances. These distances between points would not change at all if you rotated the map clockwise or counterclockwise by any amount or if you flipped it horizontally or vertically on its axes.

Fig. 1. Point map for a hypothetical 89 statement example.

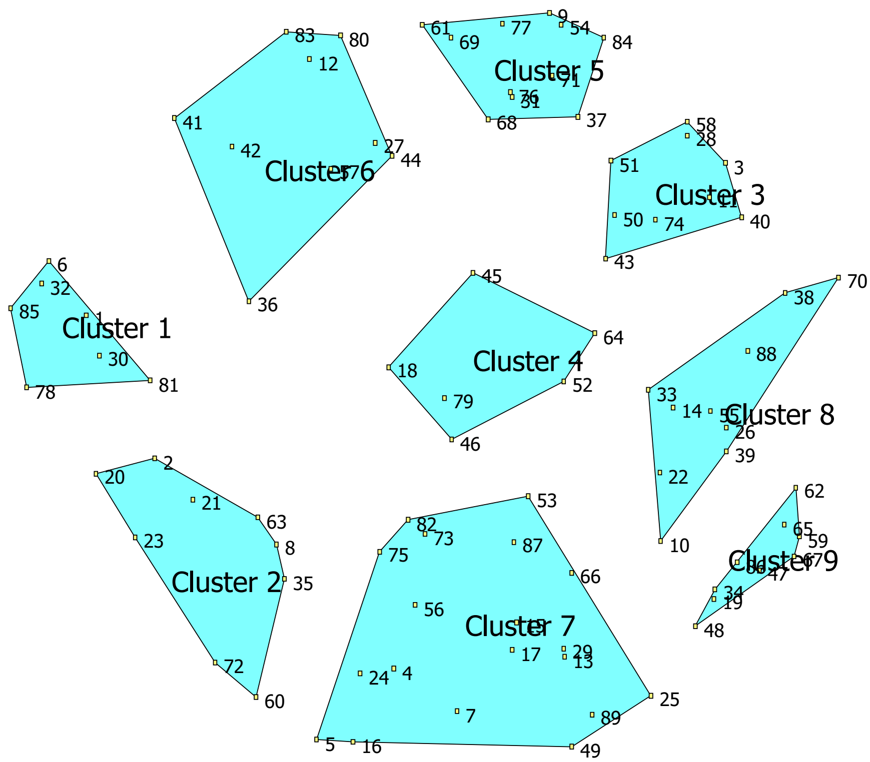

Fig. 2. A hypothetical 9-cluster map of the 89 statement point map.

The other feature worth noting is that distance between points (i.e., statements) on the map is meaningful. For instance at the top center of the map we see that statement 61 is located next to statement 69. Their proximity results from the fact that MDS placed them near each other because more people sorted them together in piles (regardless of what other statements any individual also placed in the same pile). We would expect that these statements are similar in meaning to each other as judged by the participants. The opposite is also true; statements that are farther apart are less related and are likely to represent different concepts. For instance, statement 16 on the bottom of the map is likely to be very distinguishable from statements 61 and 69.

A "cluster map" is also generated that displays the original statement points enclosed by polygon-shaped boundaries for the clusters as illustrated in Fig. 2. Because the cluster analysis uses as input the x,y coordinate matrix from the MDS analysis it necessarily results in a partitioning of the map into non-overlapping clusters. The example in the figure shows the hypothetical nine cluster solution. The shapes of the polygons are arbitrary and are constructed by connecting the outer points in each cluster. In general, the larger a cluster the "broader" the meaning of the contents of the cluster and the tighter the cluster the narrower is its meaning. For example, Cluster 6 in the figure is likely a broader conceptual one and Cluster 9 a narrower one.

The rating data are averaged across persons for each item and each cluster. This rating information is depicted graphically in a "point rating map" showing the original point map with the average rating per item displayed as vertical columns in the third dimension (Fig. 3) and, in a "cluster rating map" that shows the cluster average rating using the third dimension (Fig. 4).

It's worth noting that the point rating map shows the average rating across all participants and the cluster rating map shows the averages of those ratings across all statements in the cluster. In general, the higher an average rating, the more important it was judged to be by the participants. So, Statement 80 at the top of Fig. 3 is one of the more important statements whereas Statement 70 on the right edge of the map is one of the lowest. Similarly, in Fig. 4, Cluster 6 is one of the more important clusters overall whereas Clusters 3, 4, 8 and 9 are relatively less important.

Once the basic map structure is determined it is possible to construct any number of pattern match graphs that either compare two ratings (for all participants or any subgroups) or two groups (for any rating). Groups are determined from the demographic data that was collected. A pattern match (also known in the field of data visualization as a parallel coordinates graph (Inselberg, 2009)) is a useful visual device for showing relationships and especially for highlighting the degree of relationship between the entities being displayed. A hypothetical pattern match comparing the importance ratings for two groups is shown in Fig. 5. To create a pattern match, a vertical axis is constructed for each variable or group and set side-by-side in parallel. In the example each axis represents a distinct group of participants. The minimum and maximum values for each axis can be set to the same values (an "absolute" match) or to the highest and lowest for each variable (a "relative" match). Here relative matching is used because it allows us to see better the relative comparison. Cluster labels are shown in rank order from top to bottom along each axis, based on the average cluster value for that variable or group. A line is drawn between the axes for each cluster showing its average value on each variable or group. A pattern match quickly shows whether there are clusters that are relatively different (indicated by both rank order and slope of the line). For instance, Cluster 6 is highest on importance but is fifth from the bottom (and relatively low) in rank on feasibility. The standard Pearson product-moment correlation (r) is displayed below the graph and shows the overall relationship between the two variables or groups.

Another useful graphic for looking at the relationships between ratings with respect to the concept map is what is termed a "go-zone" plot. A hypothetical example is shown in Fig. 6. A "Go-Zone" plot is a type of bivariate graph that shows the relationship

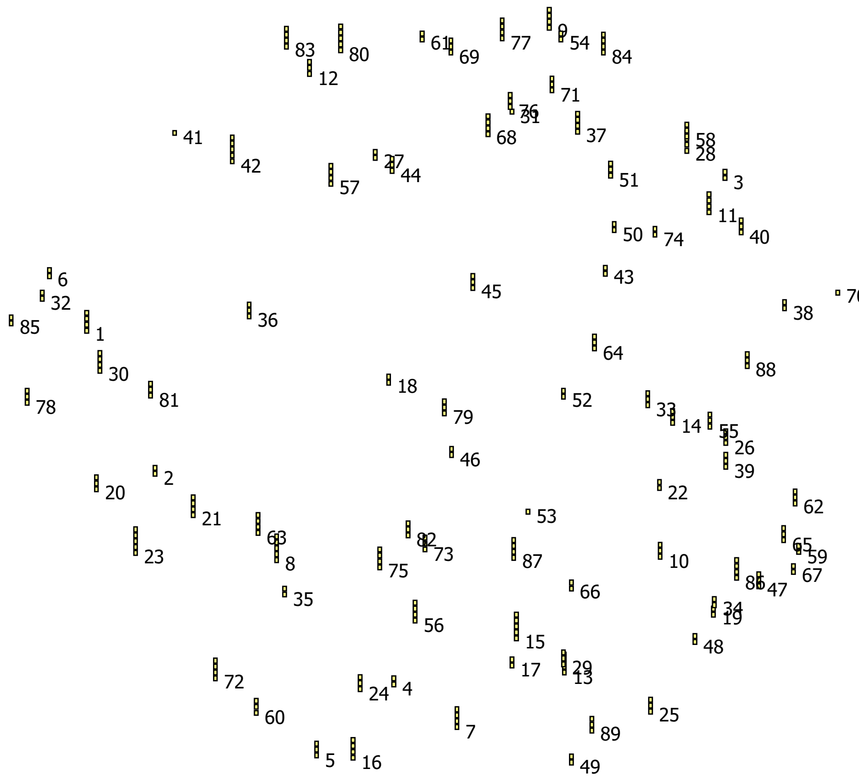

Fig. 3. Point Rating Map showing average importance of each of the 89 statements on a hypothetical map.

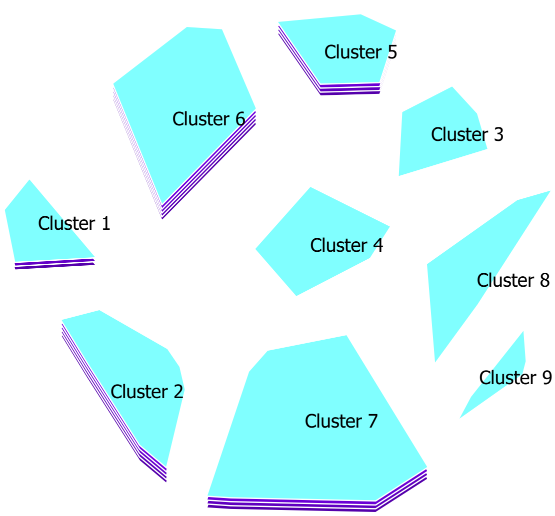

Fig. 4. Cluster Rating Map for nine-cluster solution in hypothetical concept map.

between two variables, in this case between the importance ratings for two groups. Usually a go-zone plot is shown separately for each cluster and enables one to dive into the content for that cluster and determine relative differences among statements. In a Go-Zone, each point on the plot represents the average rating for each group for each statement in the cluster. The maximum and minimum value for each axis is the maximum and minimum average for the variable across the *entire* map. So, for a single concept map, all go-zone plots will have the same maximum and minimum values. The four quadrants in the go-zone are formed by drawing a vertical and horizontal line at the cluster average of the importance and feasibility ratings. Therefore, the size and locations of the quadrants vary from cluster to cluster. The upper-right cluster is traditionally referred to as the "go-zone" because it shows the statements that are above average on both variables for the cluster. The standard Pearson product-moment correlation (r) shows the overall relationship between the two groups for the cluster. Statements in the go-zone (upper right quadrant) in Fig. 6 are by definition above average in importance for the two groups. For example, statement 15 is high in importance for both and is located in the upper right of that quadrant. Statements in the lower left quadrant are below average for both groups. Statements in the off-diagonal quadrants are above average in importance for one group and below average for the other.

Go-zone plots are particularly valuable in (and were initially created for) mapping projects where you measure both an importance and feasibility rating and wish to contrast those for all or a subgroup of participants. In this special subcase, the go-zone in the upper right would represent the statements in a cluster that were above average in both importance and feasibility and that, consequently, might be the highest priority statements to begin taking action on. These "go-to" statements make up the "go-zone" from which the plot takes its name.

CM emphasizes the value of visual products like the ones shown here and the analysis focuses on their production. However, the analysis also typically produces tabular reports of the results to augment the interpretation of the graphics.

3.5. Interpretation

The interpretation step in CM emphasizes the collaborative and participatory nature of the method. Unlike multivariate analyses in typical social research where the analyst does the interpretation, in CM a joint participatory collaborative interpretation is normative. This interpretation session follows a structured process described

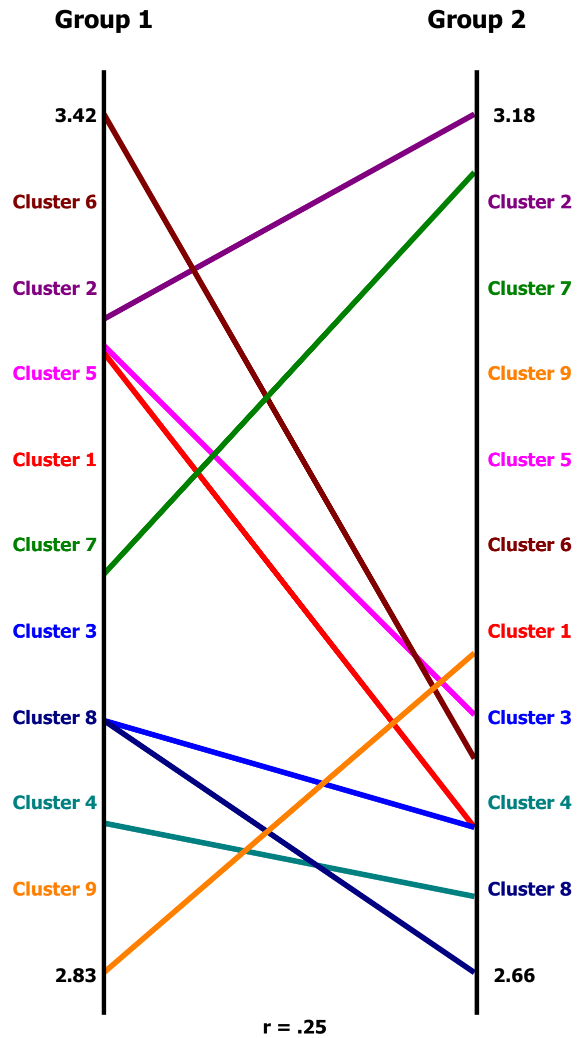

Fig. 5. Pattern Match between the cluster-level average importance ratings of two groups in hypothetical map.

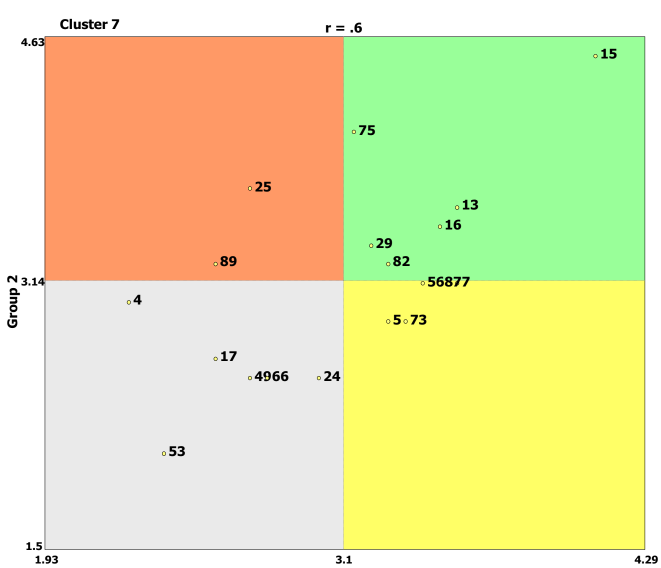

Fig. 6. Go-Zone Plot for a cluster from hypothetical concept map.

in detail in Trochim (1989b). The facilitator begins the session by showing the participants the listing of clustered statements and reminding them of the brainstorming, sorting and rating tasks performed earlier. Each participant is asked to read through the set of statements in each cluster and generate a short phrase or word to describe or label the set of statements as a cluster. The facilitator leads the group in a discussion working cluster-by-cluster to achieve group consensus on an acceptable label for each cluster. In most cases, when persons suggest labels for a specific cluster, the group readily comes to a consensus. Where the group has difficulty achieving a consensus, the facilitator suggests hybrid names that combine key terms or phrases from several individuals' labels.

Once the clusters are labeled, the group is shown the point map (Fig. 1) and told that the analysis placed the statements on the map so that statements frequently sorted together are generally closer to each other on the map than statements infrequently sorted together. To reinforce the notion that the analysis placed the statements sensibly, participants are taken on a “tour” of the map by the facilitator who identifies statements in various places on the map and examines the contents of those statements. After becoming familiar with the numbered point map, the participants are told that the analysis also organized the points (i.e., statements) into groups as shown on the list of clustered statements they have already labeled. The cluster map (Fig. 2) is shown and participants told that it is simply a visual portrayal of the cluster list. The agreed-upon cluster labels are shown on the final displayed map.

Participants examine this labeled cluster map to see whether it makes sense to them. The facilitator reminds them that in general, clusters closer together on the map should be conceptually more similar than clusters farther apart and asks them to assess whether this seems to be true or not. Participants are asked to think of a geographic map, and “take a trip” across the map reading each cluster in turn to see whether or not the visual structure seems sensible. They are asked to identify any interpretable groups of clusters or “regions." These are discussed and labeled on the map. Just as in labeling the clusters, the group arrives at a consensus label for each of the identified regions.

The facilitator notes that all of the material presented to this point uses only the sorting data. The results of the rating task are then presented through the point rating (Fig. 3) and cluster rating (Fig. 4) maps. It is explained that the height of a point or cluster represents the average rating for that statement or cluster of statements. Again, participants are encouraged to examine these maps to determine whether they make intuitive sense and to discuss what the maps might imply about the focus issue. The facilitator then presents a series of pattern matching graphs so that participants can explore the degree of consensus between different subgroups and ratings. The groups discuss each match with key points noted and recorded.

3.6. Utilization

The utilization step is an ongoing process that is very much determined by the purpose of the mapping project. If the mapping was done as part of a strategic planning exercise, the maps might be used to form workgroups for action planning. If the mapping was undertaken to accomplish operational planning such as the planning of a program, course or curriculum, the results can be used to begin to operationalize those components. In some cases the mapping was done to develop a measure or scale and the results might be adapted into a first draft of one or more instruments that can be pilot tested. In other contexts the mapping is done to generate a framework for an evaluation and the map results might be used to create a logic model or theory of change.

The go-zone plots are especially useful for 'drilling down' into the content of map results in order to begin translating a map into actionable activities and, consequently, these graphs are often integrated into the utilization process along with various reports of results that list statements by cluster or in order by average rating.

This brief overview of the steps in the method should provide enough of an introduction to the novice to enable one to understand the discussions throughout this special issue. However, the concept mapping literature is replete with hundreds of examples of applications of the methodology in real-world settings and the reader is encouraged to explore a number of these (some of which are included in this special issue) to get a better sense of how the methodology actually unfolds in practice.

4. Overview of the special issue

This special issue is designed to provide the reader with both an in-depth view of the history and practice of concept mapping and an introduction to some of the cutting-edge issues that lead into the future. The issue has been divided into sections and the articles grouped according to general common themes. However, many articles address a multiplicity of issues that the thematic classification may not adequately convey. The reader is encouraged to examine the work in this issue in detail to get the full impact of its breadth and depth.

4.1. History of concept mapping

The first section of the volume describes the history of the method's development. The article by Trochim this issue (this issue 02) is a subjective reflection on the origins and early evolution of the method by the person who started this line of work. He recalls how the method emerged from his teaching and advising work with graduate students and discusses how the immediate needs of the research they were working on and the issues that were prominent in the early 1980s interacted to influence the key developments in the approach. It's clear from this piece that concept mapping was in many ways a product of its time that incorporated ideas from participatory evaluation, multivariate statistics, program theory and theory-driven approaches, the theory of validity in experimental and quasi-experimental research (and especially construct validity), the emerging field of cognitive psychology, and the new technology of the microcomputer. He shows that the method has had a significant effect on a broad range of fields and that there are communities of practitioners in a surprisingly diverse range of universities and countries. In recounting this personal history, this article provides a good orientation to the method and many of the issues that are still affecting its evolution.

The article by Donnelly (this issue 03) focuses in even greater depth on the role that graduate education played in shaping concept mapping and how the method in turn shaped the careers of a large number of graduate students. He conducted a research synthesis of over a hundred dissertations on the method, summarizing each on 77 different variables. He points out that the method was especially suited for dissertation work because such efforts are usually among the first in a young scholar's career and typically require the articulation of a conceptual framework to guide so ambitious a research project. As one might expect, the vast majority of these dissertations focused on a topic that was previously unstudied or understudied, simultaneously the chief challenge for many graduate students and the primary appeal of concept mapping. Donnelly's synthesis confirmed that dissertation-level research has historically been and is likely to continue to be a major context for concept mapping work.

4.2. Concept mapping in context

The second section of articles in the volume addresses the relationship of concept mapping to the context in which it is applied. When speaking with individuals interested in learning about and then applying concept mapping, a recurring theme in their questions concerns the flexibility of the method. Descriptions of concept mapping in the literature may portray to the novice a method that has specific requirements for each step. In reality, the answer to the question about flexibility is that the method can be tailored to the specific needs of the project at hand. In another synthesis of the literature, Anderson and Slonim (this issue 04) reviewed the variations of concept mapping in seven articles concerned with public health. Notable here is that they describe how each of the studies addressed the aims of the study by applying variations to a step in the overall method.

The Netherlands has been one of the centers of practice for concept mapping in the world. Nabitz et al. (this issue 05) provide an excellent historical overview of the use of concept mapping in the Netherlands and, through them, in other parts of Europe. They describe ninety concept mapping projects conducted in the general area of mental health services and classify the projects into five time periods and five broad types of mapping projects: theory and model building; policy and management; planning and evaluation; quality; and, research. They present more detailed examples of three representative cases that illustrate the variety of applications and the impact of concept mapping in different contexts. In another article that reviewed 12 Dutch concept mapping projects that shared an emphasis on integrating research and practice in public health, (Bon-Martens et al., this issue 06) the authors make the case that the method was useful for integrating knowledge from science and practice, advanced the development of theory, and was useful for practical decision-making.

The paper by Vaughn et al. (this issue 7) describes the varieties of ways the concept mapping methodology has been integrated into community-engaged and participatory studies. They conducted a comprehensive review of the literature on the use of the method in such contexts and describe how the method was integrated in 103 separate publications, the major outcomes addressed and benefits associated with its use, and the methodological challenges that were identified. They distinguish between low and high levels of community engagement and make the case that concept mapping has greater impact in high engagement contexts where stakeholders are more active collaborators. They argue for greater integration of the methodology with team building, group dynamics and leadership skills as a way to encourage a more powerful and useful result.

4.3. Integrative concept mapping

The third section of this special issue describes how concept mapping can be integrated with other approaches. The paper by Stoyanov et al. (this issue 8) suggests that cognitive style is so critically important in determining concept mapping process that addressing it should be an integral task in the standard methodology itself. That is, future versions of the methodology should integrate the assessment and use of cognitive style, not just as an additional external demographic variable, but as part of the structured conceptualization process. Whether their preliminary approach to cognitive style or some variation ultimately survives as a new and integral part of the methodology remains to be seen, but their attention to issues of cognition in relation to concept mapping is a much-needed addition to the literature. Of special importance methodologically is their use of Procrustes analysis as a method of comparing two maps based on the same set of statements, in this case comparing maps of two groups with different cognitive styles. This approach could be used in any comparison of multiple maps based on a common statement set and warrants further attention in the future.

There have been only a few articles that illustrate the use of concept mapping to create a conceptual framework as the basis for scale development (see for example, Rosas & Camphausen, 2007). The article by Soellner, Lenartz & Rudinger (09 this issue) adds to and extends this literature, the authors use concept mapping to articulate the multifaceted nature of the construct of health literacy and to create an instrument to measure this construct. In an interesting variation specific to the method of concept mapping, these authors use a three-dimension solution from MDS. As pointed out earlier, ease of interpretation is a consideration that generally necessitates a two dimension solution. In this case, achieving the aims of the project did not limit the authors to two dimensions and they go on to create a qualitative model of health literacy from this map. Subsequently, they illustrate use of the items from the map to create a health literacy assessment that was subsequently tested and refined using structural equation modeling.

Systems modeling is a method that, like concept mapping, addresses complex issues, relies on multiple perspectives and visualizes results in way that is accessible to stakeholders. Lich, Urban, Dave, & Chall (this issue, 10) continue the theme using concept mapping to create a conceptual framework for the integration of other methods. In this case the authors apply concept mapping to identify program features that support youth with severe emotional disturbance making the transition to adulthood. That framework along with other sources of data served as the basis for modeling the relationships between elements in a system to support transition in a system's model. In addition to describing these relationships qualitatively, the authors created a dynamic model that allows stakeholders to test the effects of decisions in a simulated environment. This article illustrates how each method complements the others and when integrated, extends understanding and allows stakeholders to hypothesize and test actions.

4.4. Methodological advances in concept mapping

The next section of this special issue addresses methodological advances in concept mapping. One of the major uses for concept mapping, and one that is integrally linked to its central analytic method of multidimensional scaling and its historical association with construct validity, is its use in the development of new measures. The article by Rosas and Ridings (this issue 11) reviewed 23 studies where concept mapping was used for measurement development. They describe the different purposes, focus statements, participants, methods of statement generation and structuring, ways of managing the interpretation, and approaches to developing and testing the quality of measures that resulted. They conclude that concept mapping provides a good approach for establishing content validity, produced a good framework for researcher decision-making, enabled the measures' target populations to provide critical a priori input, and provided a platform for analytical and interpretive use of psychometric results. They also point to several important limitations: lack of standards for reporting concept mapping process and results; lack of information about adaptation or tailoring of the method; lack of descriptions of the relationship between the concept map and the instrument that results from it; lack of information about the subsequent use of the measures that resulted.

Hierarchical cluster analysis using Ward's method to create multiple cluster solutions followed by an interpretive task to choose the appropriate solution has been the mainstay of most concept mapping projects. There are many more approaches to clustering and Orsi (this issue 12) systematically explores how solutions differ among several types of clustering methods. Particularly noteworthy is the fact that the author goes on to illustrate how to inform the choice of the final cluster solution by applying several indices to quantitatively supplement the interpretive task.

One of the major barriers to wider use of concept mapping is integrally related to its strength – it is a method that at its heart is a complex integration of several multivariate statistical analyses. Consequently, the method either requires fairly high proficiency with and access to sophisticated statistical software or the use of a user-friendly, tailored, proprietary and for many, costly software package to make use of the approach. Bar and Mentch (this issue 13) attempted to address this issue by developing a web-based application of the method in the R language that will enable wider accessibility to the concept mapping method at virtually no cost to the user. Because R is a framework that is widely used by statisticians and scientists around the world it also has the latest and most sophisticated analytic and graphic tools available. The program introduced here makes it possible for relatively inexperienced users to run a complete concept mapping analysis and for advanced users to extend the traditional analysis by incorporating state-of-the-art statistical advances, potentially broadening and strengthening the mathematical foundation of the approach including the use of different distance metrics (maximum, Minkowski, Euclidean, city block, Canberra), options for hierarchical clustering (Ward's algorithm, single linkage, complete linkage), and Analysis of Variance (ANOVA) for statistically testing for differences between clusters or groups. The sophisticated graphics in R allow for extensive new data displays of maps and the statistical results associated with them. While the program is limited to situations where concept mapping brainstorming and structuring data are collected manually or through use of some other technology, the potential for expanding the scope and sophistication of the statistical analysis of concept mapping data is enormous.

4.5. The future of concept mapping

The special issue concludes with a look at the future of the method by McLinden (this issue 14). One of the most important events that occurred over the same span of time as concept mapping has evolved is the development of the internet. While the web has already had important implications for concept mapping, the author points out that its ultimate impacts may still lie ahead. He suggests that the global and interactive nature of the internet allows for potentially transformative involvement of participants who can collaborate on structured conceptualizations through the recently discussed approaches associated with the "wisdom of crowds", "wise crowds and wicked problems". He also discusses how the revolution in data visualization enables novel ways to represent program theory and suggests that concept mapping could have an increasingly important role in such visual development. The vision of the future he sketches for concept mapping is both a stunning look at its potential evolution and a challenge to those of us who might play a role in its realization.

Acknowledgements

The editors wish to acknowledge the extraordinary contributions of Tiffany Williams a graduate student at Cornell University, who played a critical role in helping coordinate the production of this volume, and Melida D. Busch, MLIS, Manager, Edward L. Pratt Medical Research Library, Cincinnati Children's Hospital Medical Center who provided assistance with bibliometric analysis. Finally, the editors wish to extend our deep gratitude to all of the authors

of articles in this issue who endured the vicissitudes of developing such an ambitious work and the inevitable delays that accompanied them.

References

Anderberg, M. R. (1973). *Cluster analysis for applications*. New York: NY: Academic Press, Inc.

Concept Systems Incorporated (2005). *The concept system*. Ithaca: NY: Concept Systems Incorporated [available at http://www.conceptsystems.com). (Version 4.0].

Coxon, A. P. M. (1999). *Sorting data: Collection and analysis*. Thousand Oaks: CA: Sage.

Everitt, B. (1980). *Cluster analysis*, 2nd ed. New York: NY: Halsted Press, A Division of John Wiley and Sons.

Inselberg, A. (2009). *Parallel coordinates: Visual multidimensional geometry and its applications*. London: Springer.

Jackson, K. M., & Trochim, W. M. K. (2002). Concept mapping as an alternative approach for the analysis of open-ended survey responses. *Organizational Research Methods*, *5*(4), 307–336. http://dx.doi.org/10.1177/109442802237114.

Janis, I. (1982). *Groupthink: Psychological studies of policy decisions and fiascos*, 2nd ed. Boston: Houghton Mifflin.

Kane, M., & Trochim, W. (2006). *Concept mapping for planning and evaluation*. Thousand Oaks: CA: Sage Publications.

Kreuger, R. A. (1988). *Focus groups: A practical guide for applied research*. .

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology*, 2nd ed. Thousand Oaks, CA: Sage.

Krueger, R. A., & Casey, M. A. (2000). *Focus groups: A practical guide for applied research*, 3rd ed. Thousand Oaks, CA: Sage Publications.

Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling*. Beverly Hills: CA: Sage Publications.

McDougal, W. (1920). *The group mind: A kketch of the principles of collective psychology with some attempt to apply them to the interpretation of national life and character*. Cambridge: Cambridge University Press.

Morgan, D. L. (1988). *Focus groups as qualitative research*, *Vol. 16*, .

Osborn, A. F. (1948). *Your creative power*. New York: Scribner.

Rosas, S. R., & Camphausen, L. C. (2007). The use of concept mapping for scale development and validation in evaluation. *Evaluation and Program Planning*, *30* (2), 125–135. http://dx.doi.org/10.1016/j.evalprogplan.2007.01.003.

Rosenberg, S., & Kim, M. P. (1975). The method of sorting as a data gathering procedure in multivariate research. *Multivariate Behavioral Research*, *10*, 489–502.

Stewart, D. W., & Shamdasani, P. N. (1990). *Focus groups: Theory and practice*. Newbury Park: CA: Sage.

Trochim, W., & Linton, R. (1986). Conceptualization for planning and evaluation. *Evaluation And Program Planning*289–308.

Trochim, W., Donnelly, J. P., & Arora, K. (2016). *Research methods: The essential knowledge base*. Boston: Cengage.

Trochim, W. (1985). Pattern matching, validity, and conceptualization in program evaluation. *Evaluation Review*, *9*(October (5)), 575–604.

Trochim, W. (1989a). Concept mapping for evaluation and planning. *Evaluation and Program Planning*, *12*(1) .

Trochim, W. (1989b). An introduction to concept mapping for planning and evaluation. *Evaluation and Program Planning*, *12*(1), 1–16. http://dx.doi.org/10.1016/0149-7189(89)90016-5.

Trochim, W. (1989c). Secial issue: Concept mapping for evaluation and planning. William Trochim, guest editor. *Evaluation and Program Planning*, *12*(1) .

Weller, S. C., & Romney, A. K. (1988). *Systematic data collection*. Newbury Park: CA: Sage Publications.