Concept Mapping: Mixed Methods Data for Measuring Teacher Learning in an NSF-Funded Professional Development Program (COSEE Great Lakes)

Howard Walters, EdD
Associate Professor
College of Education
Educational Foundations
Ashland University
Ashland, Ohio

ABSTRACT

The literature of teacher continuing education reflects a substantive gap in our understanding of the link between effective instructional strategies in teacher training experiences to enhance science content knowledge, and the subsequent use of that knowledge in those teachers' classroom practices with K-12 students. In 2006, the National Science Foundation funded the Centers for Ocean Sciences Education Excellence: Great Lakes, as a program to link scientists and classroom teachers. One goal of this project was to implement a series of workshops to increase the content knowledge of teachers for select science content. This study applies a mixed methods design to capture and analyze data on teacher learning in these NSF-funded workshops. Concept mapping possesses a methodological advantage over criterion-referenced pre- and post-testing in that it does not require tight alignment between taught content and the measurement instrument. As utilized in this study, the process does produce quantitative findings which are inferable to other learning settings. The article provides a detailed description of the process with example maps to allow replication of the process.

Introduction

Much has been written regarding the reform of American education in theoretical and practical terms. The trends of such reform, past and current, and real, imagined, or proposed impacts are on varied levels. At the very least, we seem currently to be refocusing school reform on the local site and on individual classroom teachers, over against the broader, systemic initiatives of previous federal administrations. As Hollingsworth and Sockett (1994) noted 15 years ago: this refocusing is mostly away from "generalizations about context" toward a greater appreciation of the contexts of schools specifically. As this reform process has occurred—and it is now long-term as a derivative of the standards movements in content disciplines and the reinvigorated "national goals" from the previous executive branch—the role of the classroom teacher in the reform process has again centralized.

Three important trends which are moving to the forefront of school reform discussions are emerging in recent literature and merit consideration. First, the role of not-for-profit (NFP) or non-governmental-organizations (NGOs) is shifting in positionality from external to internal with respect to their spheres of influence over and on school systems (Honig, 2009; Rowan, 2002). This discussion is some-what distinct from but includes the emergence of organizations such as Teach For America to positions of national prominence and influence in the school reform movement, particularly as it relates to issues of teacher quality and licensure (Hess & Petrilli, 2009). This discussion includes the increasing role of NGOs in determining specific content for teacher initial licensure and continued licensure through professional development, such as the roles of NSTA and NCTM respectively on science and mathematics standards.

Second, in the post-critical era of educational research where issues of hegemony and marginalization increase in volume, perspectives on the political and professional voices of teachers and students dominate some circles of discussion. The value that is placed on teacher knowledge and experience—or lack of value—and the input and role of teachers in school reform activities is questioned (Cook-Sather, 2009). Kennedy (2005), in a powerful treatment of school reform and the relationship between systemic administrative levels and discrete classroom realities, identifies three clusters of school reform activity: "more important content, more intellectual engagement with that content, and universal access to knowledge" (p.6). She concludes that most of these school reform efforts have failed—including those linked to teacher professional development activities and content knowledge.

Finally, some analysts question the viability of accountability cultures to adequately address, stimulate, and enhance student learning at the classroom level. Firestone (2009) juxtaposes "accountability culture" with "learning culture"—seeing these two as not necessarily related. He identifies organizational breakdown and lack of articulation between central offices or central management (the district) and the changes that are actually made at the classroom level that influence student learning.

These calls to reform—and to "reform the reform effort" are embedded with a number of philosophical beliefs and statements common to calls from earlier decades. The substance of these reform statements seems focused on the movement of content knowledge from laboratories, research centers, and the private sector to the classroom by enhancing the content knowledge and skills of classroom teachers in a way that materially and effectively changes the practice of teaching, and thereby the essence of learning among students. An earlier, important and concise treatment of the reform vision is the report of the National Commission on Teaching and America's Future (NCTAF, 1996) organized around five recommendations which remain pertinent and critical:

1. Get serious about standards for both students and teachers.
2. Reinvent teacher preparation and professional development.
3. Fix teacher recruitment, and put qualified teachers in every classroom.

4. Encourage and reward teacher knowledge and skill.
5. Create schools that are organized for student and teacher success.

Theoretical Background for Teacher Learning

Simultaneous to the development of a reformist literature has been significant expansions of the literature of Continuing Professional Education (CPE) for teachers, (particularly the refinement and disaggregation of theories for CPE design) and the literature of Epistemology and its treatment of knowledge types. These literatures remain largely disassociated, and this gap at the philosophic level contributes to the uncommunicative and unproductive nature of "dialogue" between educational theorists, researchers, and policy-makers on the one-hand, and the practitioners of education on the other. And in turn, this deficiency in communication has birthed both researchers housed awkwardly in practice centers, and shallow, practitioner-lead, research programs.

It is noted that the core of this article to follow is essentially highly practical and methodological. The purpose in the research presented here was to develop, apply, or explore the utility of a specific measurement methodology to support a theoretical learning position espoused in historic literature for teacher learning. Thus, the theoretical literature is accurate and applicable, though in some cases more than a decade old. The methodological approach to measurement is most recent in its development and application.

It is not the purpose of this article to present in philosophical terms the scope and arguments of knowledge literature in teacher education. Select publications in this literature have clearly established, however, the necessity of infusing these ideas in more practical discussions. Fenstermacher (1994) builds a compelling case for more philosophical considerations by describing the inconsistencies of epistemic weight or evidence required to define the parameters of teacher knowledge. What is meant by "a teacher knows" when used by researchers or practitioners differs based on the evidence required to sustain arguments for proof of knowledge attainment. Fenstermacher posits "teacher knowledge/formal" against "teacher knowledge/practical" as, in part, the distinction between preceptive knowledge and knowledge-through-experience—though this summarization is simplistic. He observes that in the United States, "many members of the policy-making community are embracing a view of teacher knowledge and skill that represents a limited epistemological perspective on what teachers should know and be able to do" (p.4). It is at the feet of this policy view where Fenstermacher ensconces initiatives for certification and licensure, and curricular standards and assessments at local, state, and national levels. Continuing, he concluded that as "educational policy is grounded in weak or erroneous assumptions about the nature of knowledge, there is a high likelihood that is will fail to address the problems and aspirations of education in positive and ameliorative ways" (p.4).

Moses (2002) alluded to a similar short-sightedness leading to practice mistakes among educational researchers as well. In what seems to be a "piling on" in some publications or professional circles—the usefulness, purposes, or appropriateness of empirical research, positivistic inquiry, and "traditionalism" have been questioned, criticized, and abandoned. Such a phenomenon is driven in part by postmodernist views that knowledge is *only* individual and relative. Moses discussed the spread of such a view with the adoption of qualitative methodologies in educational research. He counters, however, that:

> quantitative and qualitative research methods ought not (and cannot) be distinguished and set in opposition to one another on the grounds that quantitative methods are inherently and exclusively positivistic and suited only for confirmation, whereas qualitative methods are inherently and exclusively interpretative and suited only for understanding. (p.2)

Fenstermacher (1994) is congruent with this observation, but would clarify that typical epistemic arguments and values within these research systems are in truth, different. Following his reasoning, it would seem shortsighted—as with Moses' short-sighted policy-makers—for those seeking to describe teacher knowledge or to plan and facilitate teachers' search for knowledge to delimit what passes for acceptable knowledge, as *knowledges* are many. As the physicist Schroedinger observed, we tend to find that for which we are looking and nothing else. Thus, as observed later in this paper, researchers seeking to define the type of knowledge teachers have or should have—tend to base their conclusions on the type of knowledge they are willing to accept and nothing else. Professional development planners, likewise, tend to develop programs which allow teachers to encounter knowledge experiences framed under a single type of knowledge. Such approaches are simplistic in meeting the professionalizing needs of individual teachers who are at varying knowledge levels of both formal, preceptual knowledge founded on empirical and propositional research (new content knowledge)—and varying on practice knowledge due to time-in-service, and the quality and degree to which they have learned from reflection on the experiences had.

Yet, as Wilson and Berne (1999) observe, based on a belief that traditional professional development has not worked, we are racing toward "new and better models." They suggest, "Our readiness to embrace these new principles may, in fact, be rooted in a desire to escape collective bad memories of drab professional development workshops rather than in sound empirical work" (p.176).

In short, from either perspective, we know very little about how teachers learn, how they use what they learn, and what they need to learn to bridge the gap between where they "are" and where "we" wish them to be. Wilson and Berne (1999) believe researchers

> need to think about the knowledge they hope teachers will acquire through these learning opportunities….Stipulating a clear set of expectations for teacher learning might enable more research on the acquisition of professional knowledge….The fact that communities, as well as individuals, acquire knowledge has implications for crafting and assessing all professional development. (p.186-187)

Darling-Hammond (1997, 1998, 1999) and Corcoran's (1995) seminal work typifies a major theoretical foci: teacher knowledge as primarily content or cognitively-based, which resonates with modern professional development goals and activities. These researchers place great value on traditional formal knowledge, consistent with an emphasis on an enhanced and specific knowledge base as *the* means to effective CPE (Darling-Hammond, 1998) and improvement of the teaching profession (Corcoran, 1995). In this model, teachers should:

1. understand subject matter deeply and flexibly, thereby facilitating this understanding among students;
2. possess a knowledge of human developmental levels to include cultural and social experience;
3. incorporate a variety of teaching strategies and possess an understanding of and for different kinds of learning; and
4. exhibit a knowledge of curricular resources and technologies (Darling-Hammond, 1998).

This conceptual or theoretical position of "what constitutes a model teacher" has elsewhere been termed *competency development* and has produced a plethora of research and rhetoric attempting to define *the* competencies of a professional teacher. These competencies, according to Boyatzis (1982) (as cited in Gonczi, 1994), may include or have included lists of tasks or behaviors performed against some objective standard, the ability to think critically or a complex formulation of knowledge, attitudes, values, and skills applied within particular contexts (Gonczi, 1994; Livneh & Livneh, 1999).

Measuring Knowledge Changes in Classroom Teachers

As the above cited literature and other authors note, the ability of a classroom teacher to infuse current and accurate science information in lessons with students is highly dependent upon that teacher's science content knowledge preparation. Significant program funding has been provided nationally by the National Science Foundation for the provision of professional development for classroom teachers for the purpose of enhancing their content knowledge. This activity is founded on a decades-long belief and commitment that classroom teachers require support to bridge the gap between science content knowledge they may have acquired much earlier in undergraduate coursework, and more recent knowledge which is being created in science laboratories now. Funded as educational research activities, these programs are required to document "programmatic accomplishment and successes." Only recently, however, has attention been placed on the rigor of these programs' accomplishments as research endeavors so as to determine the research potential for instructional and assessment approaches to the broader educational field.

Methodology

One of those NSF efforts, the *Centers for Ocean Sciences Educational Excellence: Great Lakes*, was funded as an eight-state, federal and state partnership to enhance ocean science content knowledge by classroom teachers and informal educators, and their ability to infuse that knowledge in their teaching of K-12 students. Inherent in this program design was the development of a programming and learning relationship between classroom teachers (end-users of knowledge) and research scientists (knowledge creators).

Over a two year period of programming within the five year effort, a sequence of structured measurements was taken to monitor the content knowledge of teachers at the beginning of the COSEE GL programs (front end assessment), regularly throughout the program (formative assessment), and at the end of each of three intensive, week-long professional development programs (summative assessment). The results of this monitoring will eventually be used to further track and describe the infusion practices of these classroom teachers following program participation to assess the specific content areas which most regularly impact K-12 students. At four distinct times in the instructional experience of these teachers, the researcher facilitated the preparation of concept maps using a controlled implementation procedure to collect data on the knowledge of the teachers, and the organization of that knowledge into thematic and disciplinary clusters. The use of concept maps for knowledge monitoring has emerged, as cited below, as an innovative and powerful measurement approach.

The use of concept mapping to monitor changes in content knowledge has been well-addressed in research literature (All & Huycke, 2007; Arslan, 2006; Chinnappan & Lawson, 2005; MacNeil, 2007; Uzuntiryaki & Geban, 2005). Over a sequence of concept maps created individually (Arslan, 2006; Hough, O'Rode, Terman, & Weissglass, 2007), or in small groups (All & Huycke, 2007; Hong, Losh, & Turner, 2007), changes in both the scope of content knowledge and the complexity of that knowledge can be observed and documented. Additionally, other NSF funded teacher education initiatives have used concept mapping procedures to isolate and measure growth in content knowledge of teachers (Hough et al., 2007). Novak and Gowin (1984), seminal researchers in the development of concept mapping for facilitating and assessing science education, note that concept maps were "developed specifically to tap into a science learner's cognitive structure and to externalize, for both the learner and the teacher to see, what the learner already knows" (p.40). The ability to "tap into" and visualize the content learning of classroom teachers in these workshops was viewed as an important approach to not only monitor the effectiveness of COSEE GL in implementing its programming goals, but a way to leverage the assessment toward a contribution to learning theory, measurement theory, and program assessment literature simultaneously.

Using these field-tested approaches, the evaluator and internal program team for COSEE GL have systematically implemented data collection in the workshops (week-long, intensives meeting the content requirements and time for 3 credit hour graduate courses). These data, in the form of individual and group created lists of science concepts and propositions, and individual and group created concept maps, support a conclusion that individual classroom teachers have increased their content knowledge of science in COSEE GL programs. Further, these data visualize the emergence of a *hierarchical network* (Hough et al., 2007; Novak & Gowin, 1984) of science content related to the Great Lakes that is substantially more complex and deeper at the end of program participation than at the beginning of program participation for teachers. The following narrative and charts demonstrate selections of the broad set of data that has been collected to date that support this conclusion and the method of analysis. Following this selection, analyses and findings from the broader set of concept map data are described and discussed.

Analysis of concept maps necessitates the use of a complex jargon which has emerged in the literature over time. Key vocabulary from this jargon with definitions are listed below, and have been substantially quoted from Hough et al. (2007) but are quite standard in the literature.

1. Concept—an idea/term/phrase contained in a single polygon or oval on a map;
2. Root—the central phenomena or initial definition/term/concept on a map;
3. Link—a line connecting two concepts;
4. Level—all concepts which share a distance X from the Root;
5. Depth—the distance in concepts for the longest chain of the map;
6. Width—the number of concepts in the largest level;
7. Chunk—a group of linked concepts for which the leading concept has at least two successors; and
8. Crosslink—lines connecting two separate chunks.

The following charts (one through three) represent a sequence of three concept maps created by teachers who participated in the COSEE GL Lake Huron Exploration Workshop in Alpena, MI during the summer of 2007. This workshop, implemented for participants over a seven day period, included a series of guest presentations from scientists and science educators, and included intensive field experiences nearly each day of the workshop. These three charts are selected from a larger set of maps created daily over the course of the week, both by individuals working in personal journals, and by small, ad hoc groups of students. There are three initial levels of emerging complexity evident in this smaller selection of maps provided here.

First, there is an evolving complexity in the use of vocabulary terminology. Second, there is an evolving complexity in the levels and number of linkages identified to connect the core phenomena (termed "root" in the literature), i.e. there are more connections drawn as the week progresses, and there are more interconnections drawn as the week progresses.

Chart 1 includes the overall title/content of the map (the root), which was provided by the project evaluator who led the mapping session. Extending from this root, a single connecting verb "includes" is linked to four terms encountered by participants. Further analysis of the situated context of these four terms within the workshop itself as delineated by the planned activities during day one (which further included orientation sessions and reading materials sent previously to participants) has not been completed.

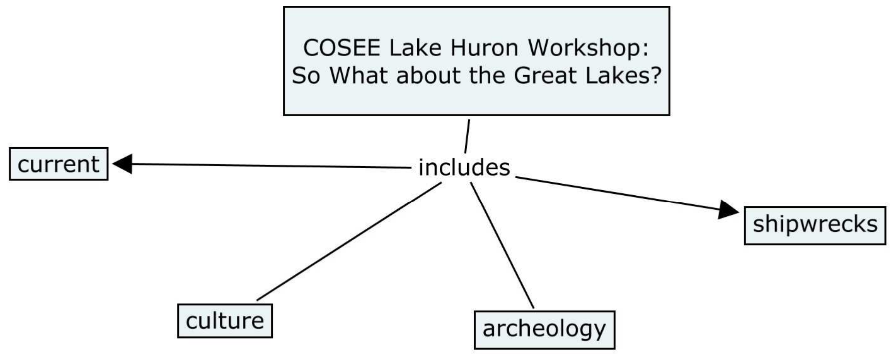

Chart 1. Lake Huron exploration workshop (2007) Monday map.

The prompt for creation of this map was intentionally vague—with the root providing the prompt. The “so what” allowed respondents to freely associate interesting phenomena to the content area, without regard for correctness and pertinence to the workshop. Additional analyses will summarize all of the maps developed at this stage of the workshop to identify the overall impact of the workshop.

Chart 2 includes the same root prompt developed for and presented on the fourth day of the workshop. The first level core phenomena that emerge for this group of respondents includes more categorical level content place-holders, i.e. Terrestrial Science instead of plants, or Aquatic Science instead of fish. Additionally, the complexity of linkages is more obvious—with a mathematical increase in linkages. Additionally, the linkages are now on two or more levels, suggesting the increase in content experiences during the workshop at this point in time (Wednesday workshop). This phenomena is termed the “depth” of the concept map, and is a primary component of complexity in a map.

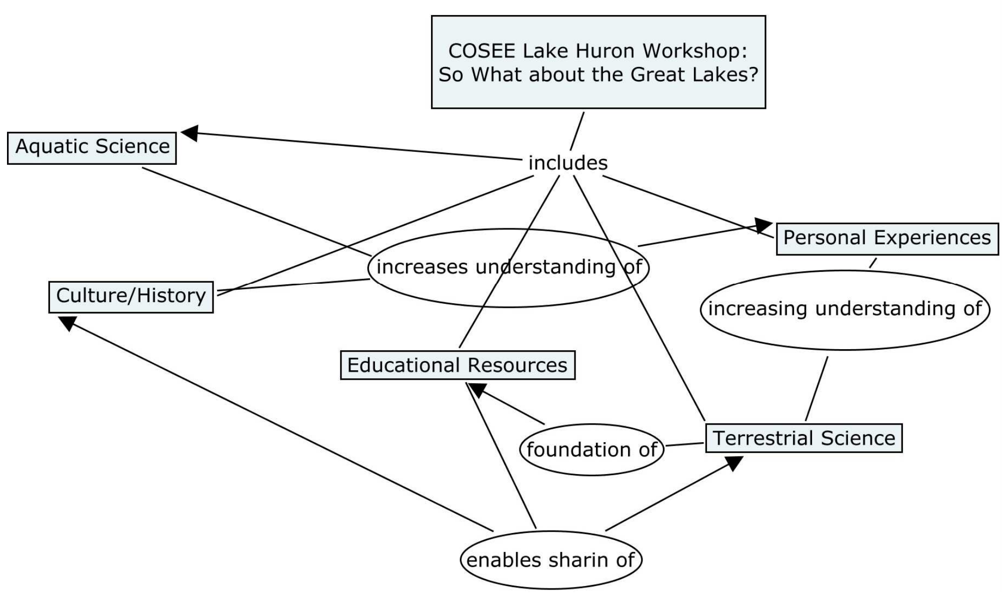

Chart 2. Lake Huron exploration workshop (2007) Wednesday map.

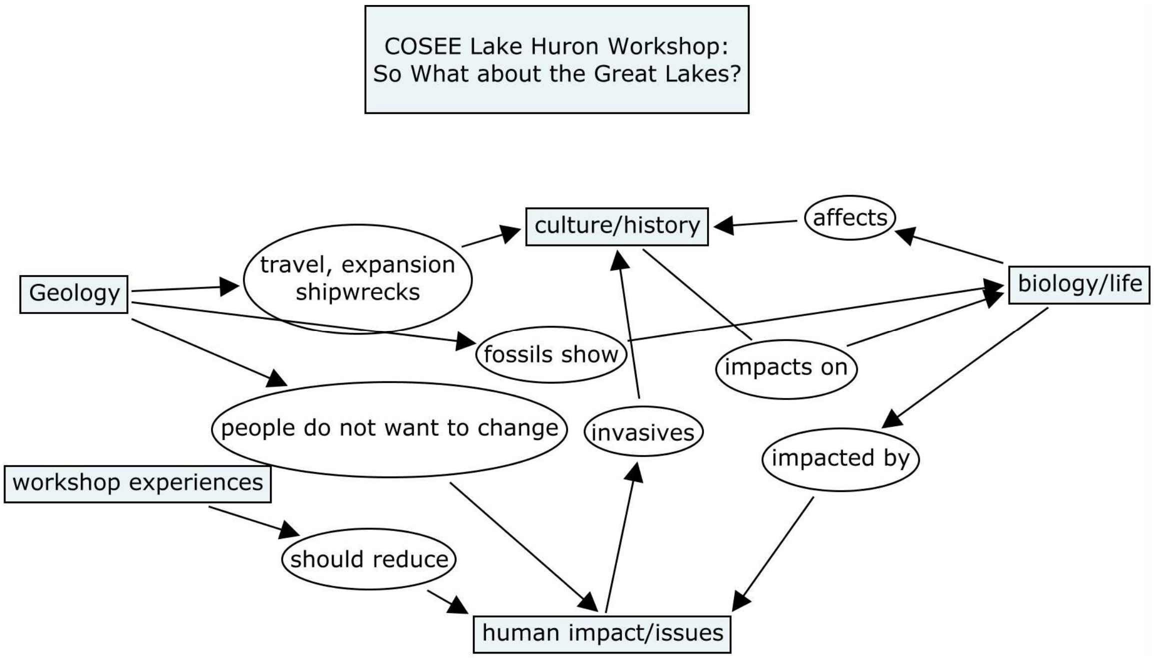

Chart 3. Lake Huron exploration workshop (2007) Friday map.

Chart 3 is the final concept map for one group collected on Friday morning of the workshop. The complexity, again, is increased from level two. The single connector "includes" linking the root to the overall map has been dropped by respondents as too prohibitive, but conceptually continues to guide the map's content boundary. The first level of concepts is linked in some cases in two directions with each other, as the lines reciprocally from biology/life to and from culture/history demonstrate. Further, multiple connections from one concept to multiple phenomena are evident, as the links from geology to culture/history, biology/life, and human impact/issues demonstrate. These examples also demonstrate both chunks and crosslinks.

For each of the above three maps, a set of values was derived through counting. These values are termed hierarchical scores (Novak & Gowin, 1984) and selectively combine to create a hierarchical structure score or HSS (Hough et al., 2007). The HSS is calculated as width (w) plus depth (d) of each map. Table 1 also includes a Weighted Crosslink following Novak and Gowin (1984, p. 107) which is 2 (crosslink score). These values are included by category and map in Table 1. Further analyses of these values follow.

Table 1

Hierarchical Scores for Charts 1, 2, and 3

Score	**Chart 1**	**Chart 2**	**Chart 3**
Concepts	5	10	13
Width	4	5	5
Depth	2	3	6
Chunks	0	5	7
Crosslinks	0	4	6
Weighted Crosslinks	0	8	12
HSS	6	8	11

Based on this data collection approach, i.e. collection of maps early in the workshop and at the end of the workshop for each of the three delineated workshops; multiple maps have been created by participants working individually and/or in small groups. A set of 20 group-created concept maps have been coded as in the example maps provided above to yield an appropriate data matrix for statistical analyses following Hough et al. (2007). These 20 maps include 10 each from the beginning of the weeklong programs and 10 from the end of the weeklong programs. Collectively, these 20 maps account for approximately 40 classroom teachers. Table 2 provides the descriptive statistics for the 20 maps collectively.

Table 2

Descriptive Statistics for Map Variables

	N	Range	Min.	Max.	Mean	Std. Deviation
CONCEPT	20	49.00	3.00	52.00	16.1500	12.66273
WIDTH	20	10.00	3.00	13.00	5.4000	2.13739
DEPTH	20	7.00	2.00	9.00	4.6000	1.87504
CHUNK	20	8.00	.00	8.00	3.7500	2.46822
CROSSLIN	20	9.00	.00	9.00	2.8500	3.13344
HSS	20	17.00	.00	17.00	7.8000	5.29747
Valid N (listwise)	20					

A quantitative analysis was conducted to determine the statistical relationship between the early workshop maps and the end of workshop maps collected on the groups of teachers (Table 3). Specifically, a MANOVA was calculated using as dependent variables the number of concepts (CONCEPT), the Hierarchical Structural Score (HSS) for each map (calculated as W+D), the number of chunks in each map, and the number of Crosslinks in each map. It is noted that Hough et al. (2007) used the raw scores for crosslinks for maps in their analyses, whereas Novak and Gowin (1984) recommended using a weighted crosslink score of 2 or 3 times the crosslink number, based on an observation that the crosslink was a more substantive indicator of conceptual complexity than the width or depth of concepts alone or summed (the HSS). This researcher has adopted Novak and Gowin's original recommendations in this analysis, using the weighted crosslink (WCL).

Table 3

Tests of Between-Subjects Effects

Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
CONCEPT	1394.450(b)	1	1394.450	15.193	.001	.458
CHUNK	76.050(c)	1	76.050	34.481	.000	.657
WCL	84.050(d)	1	84.050	14.760	.001	.451
HSS	369.800(e)	1	369.800	40.737	.000	.694

Overall, it was found that there was a statistically, significant increase in content knowledge and knowledge complexity between the two groups of maps. Consequently, post hoc testing was calculated. Each of the individual dependent variables were found to be statistically, significantly greater for the end of workshop maps over the beginning of workshop maps using a Sidak correction to obtain more conservative results in the post hoc testing. The eta-values are a measure of the effect size of each of the individual dependent variables, and are strong. These data support a conclusion that there was a strong, positive increase in both content knowledge (number of concepts) and structural complexity (WCL and HSS) of that content knowledge for these classroom teachers. The eta score for HSS, the main unit of complexity of content knowledge, reveals that 69% of the overall score change is associated with increased complexity of knowledge by the teachers. The number of concepts accounted for approximately 45% of the overall change in dependent scores—significant but less than the complexity score. Levine's Test finds the internal variances for CONCEPT ($p<.05$) and WCL ($p<.05$) at a cautionary level, but supports the use of the HSS and CHUNK scores (not significant).

Pairwise comparisons of the individual dependent variables were calculated using ANOVAs using the pre- and post-maps as the independent variables. Tables 4 and 5 support a conclusion that each dependent variable measured on the post-program maps was statistically greater than that variable measured on the pre-program maps. Correcting for multiple, post hoc inferences using Sidak, each MAP 1 to MAP 2 comparison was statistically significant in one-tailed testing, only considering positive growth in the post-measures as important or desired.

In summary, based on the individual number of concepts included on the group created maps, there was a significant increase in the scope of science content knowledge associated with the Great Lakes by classroom teachers at the two, Lake Exploration workshops (land-based) and the one, Lake Guardian (ship-based) workshop from which map data were compiled for the analyses. Second, using the overall width and depth of the maps based on the concepts and the number and structure of their linkages as proxies for hierarchical structural complexity, there was a statistically significant improvement in the complexity of understanding of Great Lakes science content by these classroom teachers across these three workshops. This finding supports a conclusion that the COSEE GL team is effectively reaching one of its important program objectives. Finally, these programs were facilitated by science educators incorporating research scientists as the primary instructors who conveyed the new science content knowledge to these teachers. Consequently, these statistical analyses support a finding that these scientists were effectively meeting their "broader impact" objectives and concerns, and were effective in bridging the gap between their research findings and the cognitive needs of these classroom teachers.

Table 4

Univariate Estimates of Map Score Differences

Dependent Variable	MAP	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
CONCEPT	1	7.800	3.030	1.435	14.165
	2	24.500	3.030	18.135	30.865
CHUNK	1	1.800	.470	.813	2.787
	2	5.700	.470	4.713	6.687
CROSSLIN	1	1.600	1.509	-1.571	4.771
	2	9.800	1.509	6.629	12.971
HSS	1	3.500	.953	1.498	5.502
	2	12.100	.953	10.098	14.102

Table 5

Significance of Pairwise Comparisons

Dependent Variable	(I) MAP	(J) MAP	Mean Difference (I-J)	Std. Error	Sig.(a)
CONCEPT	1	2	-16.700(*)	4.284	.001
	2	1	16.700(*)	4.284	.001
CHUNK	1	2	-3.900(*)	.664	.000
	2	1	3.900(*)	.664	.000
CROSSLIN	1	2	-8.200(*)	2.134	.001
	2	1	8.200(*)	2.134	.001
HSS	1	2	-8.600(*)	1.347	.000
	2	1	8.600(*)	1.347	.000

Based on estimated marginal means
* The mean difference is significant at the .05 level.
a Adjustment for multiple comparisons: Sidak.

Conclusions and Discussion

It is clear from the analyses that the teachers who participated in the series of workshops in this project have benefitted from significant increases in their content knowledge, and from a significant increase in their ability to structure that knowledge into meaningful and complex patterns or sets of content. It will be important to track these teachers as they infuse the content into authentic classroom settings, to describe the nature, scope, and characteristics of their use of the content in teaching before useful conclusions regarding the longer-term impacts of this program can be reached. However, it seems clear that several important observations can be made at this time.

First, literature and program reports alike are replete with program assessment data that are limited to participant-reported responses of the quality and appropriateness of programs. These data are substantially limited because of self-report bias and the nature of teacher responses to professional development generally. For example, we know that professional development experiences that are linked to stipends, free materials and supplies, and travel or other exotic program characteristics—or winsome and personable staff—will obtain more positive feedback from participants. We have little evidence that these positive emotional responses translate into better teaching, or in the short-run, whether they even translate into effective cognitive learning for participants initially. Directly measuring changes in content knowledge, from a baseline at the beginning of a learning event, at multiple points across that event, and at the end, allows researchers and program coordinators a more objective vantage point to view program impact.

Second, the results of this study are marked by a singularly important methodological improvement and advantage to the task of directly measuring content change among program participants. Historically, the use of a criterion referenced pre-test and post-test has provided an important and nearly irreplaceable method for this measurement. Unfortunately, this requires advanced awareness of program content, and once that content has been incorporated into an instrument, it is difficult to "change on the fly." Additionally, a somewhat advanced knowledge of test construction is necessary to create valid instruments. The use of the concept map avoids each of these problems. No advanced awareness of the specific content items to be taught—beyond a general awareness of the broader themes or disciplines of the learning event—is necessary. In the case of this research effort, the researcher began, literally, with a blank piece of paper and a knowledge that the content for the week-long program was generally within the framework of the Ocean Literacy Principles. Nevertheless, by using multiple maps developed over time and a well-reviewed analysis procedure to convert the narrative responses to mathematical data, the researchers are able to develop a significance level for the change in content knowledge and knowledge complexity that emerged across the program. This quantitative value, in turn, will allow the researchers to compare and infer findings across programs for theory development.

Finally, significant investment into teacher professional development is driven by the hope that the content provided in these workshops is ultimately infused in classroom teaching. Nevertheless, little empirical work has been done to track this. We rarely are informed whether the content provided teachers ever makes its way to their students. The concept maps which have been developed as primary data for this current manuscript are being maintained for secondary analysis. Currently, data are being collected at 6-month, 12-month, and 18-month intervals on the teachers who have participated in these workshops. They are being asked to provide lesson plans and sample student activities. These materials will be summarized and then compared back to the concept maps to identify which science content was ultimately taught to students, at what grade levels, in what format and over what time period. This will allow the project management and the researcher to describe the relationship between the original workshop activities, i.e. were the lab based, lecture based, ship based, field based, or other; did these activities result in enhanced content as measured on the concept maps reported in this manuscript; and, did that content eventually appear in classroom instruction of students.

References

All, A. C., & Huycke, L. I. (2007). Serial concept maps: Tools for concept analysis. *Journal of Nursing Education*, *46*(5).

Arslan, M. (2006). The influence of teaching note-taking and information mapping on learning and recalling science. *The Turkish Journal of Educational Technology (TOJET)*, *5*(2).

Chinnappan, M., & Lawson, M. J. (2005). A framework for analysis of teachers' geometric content knowledge and geometric knowledge for teaching. *Journal of Mathematics Teacher Education*, *8*, 197-221.

Cook-Sather, A. (2009). Translation: An alternative framework for conceptualizing and supporting school reform efforts. *Educational Theory*, *59*, 2.

Corcoran, T. B. (1995). *Transforming professional development for teachers: A guide for state policymakers.* Washington DC: National Governors' Association through a grant from the Carnegie Corporation of New York.

Darling-Hammond, L. (1997). What matters most: 21st century teaching. *Education Digest, 63*(3), 4-9.

Darling-Hammond, L. (1998). Teacher learning that supports student learning. *Educational Leadership, 55*(5), 6-11.

Darling-Hammond, L., & Sykes, G. (Eds.) (1999). *Teaching as the learning profession: Handbook of policy and practice.* San Francisco: Jossey-Bass.

Fenstermacher, G. D. (1994). The knower and the known: the nature of knowledge in research on teaching. *Review of Research in Education, 20*.

Firestone, W. A. (2009, May). Accountability nudges districts into changes in culture. *Phi Delta Kappan*, 670-676.

Gonsczi, A. (1994). Competency based assessment in the professions in Australia. *Assessment in Education, 1*(1).

Hess, F. M., & Petrilli, M. J. (2009, February/March). Wrong turn on school reform. *Policy Review*, 55-68.

Hong, G. E., Shen, S. L., & Turner, J. (2007). A review of studies on collaborative concept mapping: What have we learned about the technique and what is next? *Journal of Interactive Learning Research*, *18*(4).

Honig, M. I. (2009). "External" organizations and the politics of urban educational leadership: The case of new small autonomous schools initiatives. *Peabody Journal of Education, 84*, 394-413.

Hough, S., O'Rode, N., Terman, N., & Weissglass, J. (2007). Using concept maps to assess change in teachers' understandings of algebra: A respectful approach. *Journal of Math Teacher Education*, *10*, 23-41.

Kennedy, M. (2005). *Inside teaching: How classroom life undermines reform*. Cambridge, MA: Harvard University Press.

Livneh, C., & Livneh, H. (1999). Continuing professional education among educators: Predictors of participation in learning activities. *Adult Education Quarterly, 49*(2), 91-106.

MacNeil, M. S. (2007). Concept mapping as a means of course evaluation. *Journal of Nursing Education*, *46*(5).

Moses, M. S. (2002). The heart of the matter: Philosophy and educational research. *Review of Educational Research*, *26*.

National Commission on Teaching and America's Future (NCTAF) (1996). *What matters most: Teaching for America's future.* New York: Author.

Novak, J. D. (1998). *Learning, creating, and using knowledge: Concept maps as facilitative tools in schools and corporations*. London: Lawrence Erlbaum Associates.

Novak, J. D., & Gowin, D. B. (2006). *Learning how to learn*. New York: Cambridge University Press.

Rowan, B. (2002). The ecology of school improvement: notes on the school improvement industry in the United States. *Journal of Educational Change, 3,* 283-313.

Uzuntiryaki, E., & Geban, O. (2005). Effect of conceptual change approach accompanied with concept mapping on understanding of solution concepts. *Instructional Science*, *33*, 311-339.

Wilson, S. M., &Berne, J. (1999). Teacher learning and the acquisition of professional knowledge: An examination of research on contemporary professional development. *Review of Research in Education*, *24*.