Pergamon

Teaching & Teacher Education, Vol. 10, No. 6, pp. 589-599, 1994

0742-051X/94 S7.00+0.00

0742-051X(94)E0003-N

CONCERNING THE RELATIONSHIP BETWEEN TEACHERS' THEORETICAL ORIENTATIONS TOWARD READING AND THEIR CONCEPT MAPS

JOHN R. MERGENDOLLER and COLIN H. SACKS

Beryl Buck Institute for Education, Novato, **CA, U.S.A.**

Abstract-Recent research in teacher cognition has shown that concept maps, or ordered trees, are sensitive to differences in teacher experience and to change resulting from participation in academic courses. In the current study, we examined the usefulness of concepts maps for detecting differences in kindergarten teachers' theoretical orientation (whole language versus phonics), and we assessed the stability of their concept maps over time. Both the structure and content of teachers' concept maps were examined at three times across an academic year. Results revealed that content but not structure measures differentiated whole language from phonics teachers. Specifically, whole language teachers generated more nodes related to child affect and individual differences between children. Teachers' maps were quite stable structurally and moderately stable in terms of content. The results suggest that concept maps are useful as a measure of theoretical orientation and that both structure and content of concept maps should be considered in studies using the ordered tree technique. Directions for future research are discussed.

In recent years a number of studies have used the ordered tree technique, or concept maps, as a means of assessing teachers' cognitive structures about teaching (e.g., Beyerbach, 1988; Roehler, Duffy, Conley, Herrmann, Johnson, & Michelsen, 1990; Strahan, 1989). Although studies have varied widely, both in methods used to obtain teachers' concept maps and in scoring of these maps, results converge in finding that the maps are sensitive to differences between more and less experienced teachers, and that they are useful for measuring cognitive change resulting from participation in academic courses.

In the current study, we sought to assess the usefulness of the ordered tree technique in detecting differences in teachers' theoretical orientations toward the teaching of reading (Deford, 1985). In addition, we sought to assess the stability of teachers' experience and their ordered trees. Before discussing the current study further, we present a brief review of recent studies using the ordered tree technique.

Naveh-Benjamin, McKeachie, Lin, and Tucker (1986) examined the concept maps of undergraduate students before, during, and after a Psychology of Aging course at a major university. They found that as the course progressed, students' concept maps became more organized, had a greater average depth, and became more similar to a map generated by the teacher of the course. These trends were especially strong for those students who performed well in the course (as measured by final grade), relative to those who performed less well.

Beyerbach (1988) examined the concept maps of undergraduate teacher education students before and after courses at three different levels of the program. She found that within each course, students' maps became more similar to their instructor's, and more differentiated. Further, the content of post-maps was different from pre-maps, reflecting more emphasis on strategies and processes of teaching, as well as more attention to the process of student evaluation.

We would like to thank Yolanda Bellisimo, Marie Kanarr, and the anonymous reviewers for their contributions to this paper.

Strahan (1989) examined the concept maps of experienced and novice teachers before and after a teacher education course. He found that the maps of experienced teachers contained more concepts, chunks, links between concepts and were more organized, relative to those of less experienced teachers. Further, the course led to increases in these measures for both experienced and novice groups. Examination of the content of the maps revealed that the experienced teachers were generally more "student centered," while novices' maps were initially more "teacher oriented," becoming more student centered after the course.

In a similar study, Roehler et al. (1990) found that expert teachers' concept maps were more extensive and coherent than were novices', and that novices' maps became more like experienced teachers' maps as they proceeded through two reading methods courses. Further, these investigations present data from a limited subsample of subjects suggesting that students with more elaborate concept maps respond with more "responsive elaboration" in teaching situations where a child has failed to grasp a concept.

Taken together, these studies suggest that concept maps, or ordered trees, are useful for representing teachers' cognitive structures about teaching, and for documenting changes in those structures.

The primary aim of the current study was to see if concept maps can be used to demonstrate cognitive differences between teachers whose theoretical orientations toward the teaching of reading are known to differ. Specifically, we were interested in testing whether the ordered tree technique could be used to distinguish teachers who advocated "whole-language" (e.g., Goodman, 1986) or "phonics-based" (e.g., Chall, 1967) approaches to reading instruction.

Derived in part from the writings of Dewey (1929), Piaget (1952), and Smith (1988), the whole-language approach asserts that children learn language most effectively at their own developmental pace through exposure to language-rich environments and quality literature, rather than through formal instruction in the discrete skills identified and defined by behaviorally oriented reading researchers. The instructional implications of this approach are that teachers should teach children at their own developmental pace, maximize their exposure to learning activities emphasizing oral language use and quality literature, and provide ample extracurricular opportunities for literature-related activities.

The antithetical approach is the "skills based" or "phonics" approach, based on the tenets of behaviorism (Heald-Taylor, 1989). Seen as the more "traditional" approach, this method emphasizes the teaching of language subskills (e.g., letter recognition, sound-symbol association) with a gradual movement toward larger language units, and eventually literature.

Given the differences between these theoretical approaches to teaching reading, we predicted that the cognitive maps of teachers known to emphasize the whole language approach would contain a higher proportion of nodes concerned with individual differences between children, affect toward the learning of reading, and outside class activities associated with reading, than would the maps of more phonics-oriented teachers.

To measure teachers' theoretical orientations, we used an instrument whose reliability and validity have already been established, the Theoretical Orientation to Reading Profile (TORP; Deford, 1985). Theoretical Orientation was assessed prior to the generation of teachers' cognitive maps. Further details concerning the TORP are provided in the Method section.

Our second prediction was that teachers'maps would be relatively stable over time. During the course of the study, all teachers were participating in inservice training which included semimonthly workshops. Topics covered included "Literacy Indicators in the Environment," "Cultural Diversity in the Classroom," and other subjects relevant to early childhood education. Some workshops were philosophically consistent with the whole-language approach to reading instruction, although whole language teaching was not the major thrust of this professional development program. However, given that only five workshops took place during the course of the study, and given that similar workshops had been occurring previous to the onset of the study, we do not treat them as "intervention," but rather as an environmental factor which bears mention. We predicted little or no change in teachers' cognitive maps across the course of the study, and either no change or a slight whole-language shift in teachers' theoretical orienta-

tions, given the content of some of the workshops.

Third, we predicted, based on the findings of Beyerbach (1988), that more experienced teachers would generate a higher percentage of nodes involving concrete examples of instructional strategy. We also predicted that more experienced teachers would generate a higher percentage of nodes involving affective factors and individual differences between children, following Strahan (1989). Furthermore, we predicted deeper and better differentiated cognitive maps from more experienced teachers, following Beyerbach (1988), Roehler et al. (1990), and Strahan (1989).

Finally, we sought to examine the usefulness of a new means of coding the content of concept maps. Both Beyerbach (1988) and Strahan (1989) examined content, as well as structure, of their subjects' concept maps, but their data were limited to interviews with subject about their maps. Instead, we generated five coding categories to be applied to the maps, intended to reflect important and divergent aspects of teaching of reading.

To summarize, our predictions were as follows: (a) Teachers' theoretical orientation (TORP) scores would predict which categories were represented in their cognitive maps, but they would not predict structures of those maps; (b) high correlations across time would be found for both structure and content measures; (c)more experienced teachers would generate deeper and better differentiated cognitive maps. Further, more experienced teachers would use a higher percentage of nodes involving affective factors and individual differences between children and involving concrete examples of instructional strategy.

Method

Participants

Sixty teachers from six different school districts in the San Francisco Bay Area participated in the current study. Teaching experience ranged from 1 to 35 years (M = 17.269, SD = 9.098). Twenty-three teachers generated codable concept maps at time 1, 31 at time 2, and 33 at time 3. Twenty-nine teachers completed codable concept maps at more than one time; 14 completed codable concept maps at all three administration times.'

Measures

Theoretical Orientation Toward Reading

The Theoretical Orientation to Reading Profile (TORP) has been shown to be a reliable, valid instrument for assessing teachers' theoretical orientation to reading (Deford, 1985). This questionnaire assesses teachers' scores on three dimensions: (a) phonics, which emphasizes smaller than word level units, with a gradual progression toward words and sentences, (b) skills, which emphasizes development of a sight-word vocabulary, and (c) whole language, which emphasizes familiarizing children with quality literature, and working down toward small language units. Scoring of each dimension is independent, but skills and phonics scores are highly positively correlated, and inversely correlated with whole language scores. Deford (1985) reports that the construct validity of the TORP was established by administering the instrument to a sample of 90 teachers selected for their differing orientations (30 in each group) by area professionals who had seen them teach. The alpha reliability of the instrument was found to be .80, suggesting acceptable internal consistency. In the current study, we examined phonics, skills, and whole-language scores separately, and created a bipolar whole-language/phonics-orientation score by subtracting teachers' phonics scores from their whole-language scores.

Concept Maps: Structure and Content

Drawing on previous studies, we scored teachers' concept maps for both structure and content.

Measures of structure. Given the diversity of approaches found in previous studies, we used five separate approaches to measure cognitive structure. First, the ***number of nodes*** in each map was arrived at by counting the number of separate entries in a given tree. In cases where a "list" was attached to a given node (e.g., "adventure", "comic", "children's", "coloring" listed next to the node "books"), each member of the list was counted as a separate node. Second, ***number of links*** was arrived at by counting the number of links between nodes. (This score should be equal to or greater than ***number of nodes.)*** Third, ***number of chunks*** was arrived at

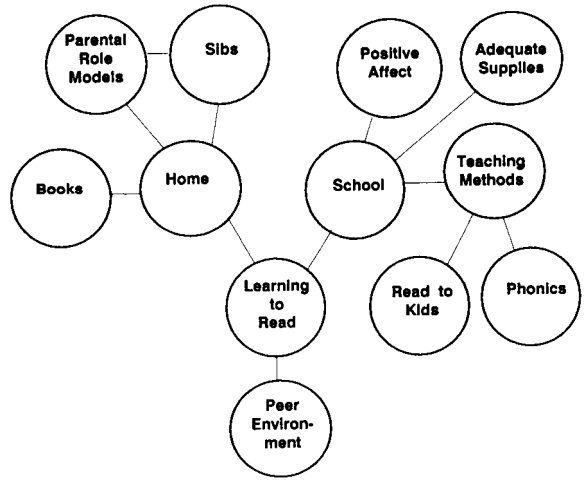

Figure 1. A sample concept map. This sample concept map has 11 *nodes,* as "learning to read" is not counted. It has 12 links, as "parental role models" and "sibs" are linked, though they need not have been. It has three *chunks,* or nodes with two or more deeper level nodes attached to them: "home," "school", and "teaching methods" are the centers of these three chunks. The maps's average *depth* is 1.33, as there are two nonterminal level 1 nodes ("home" and "school"), and one nonterminal level 2 node ("teaching methods"). Finally its *total depth* is 21, as are three level 1 nodes, six level 2 nodes, and two level 3 nodes.

by counting the number of nodes which have two or more deeper level nodes attached to them. For example, the node "books", with "adventure" and "comic" attached as deeper level nodes would constitute a chunk. However, "books" with only "adventure" attached would not. Number of chunks has been suggested by Naveh-Benjamin et al. (1986) as a measure of organization. Fourth, *average depth* was arrived at by multiplying the level of depth by the number of nonterminal nodes at that level, summing these scores for each level, and dividing the total by the number of nonterminal nodes in the map. This measure is also adapted from Naveh-Benjamin et al. (1986). As an example, a map with 5 level one nodes, 10 level two nodes (all nonterminal) and 15 terminal level three nodes, would have an average depth score of 1.67) ([l x 5 + 2 x 10]/15). Fifth, *total depth* was arrived at by multiplying the level of depth by the number of nodes at that level, and summing these scores for each level. This measure was generated because it was found that the average depth score was exaggerated for maps with few nodes and especially few level one nodes. A sample map and structural scoring are shown in Figure 1.

Measures of content. We also used five categories to typify the issues enumerated by teachers within their maps. These categories are exploratory and were chosen because they represent diverging aspects of teachers' belief systems (c.f., Clark & Peterson, 1986). First, *affective or emotional support* includes all nodes which make direct reference to the affective or emotional state of the child or parent/teacher. Examples include "leads to frustration," and "child must enjoy the task." Second, *sensitivity to individual or developmental differences* include all nodes which make reference to individual, gender, race, age, or developmental differences between children. Sample nodes include "developmentally appropriate" and "easier if child speaks English." Third, *outside class intellectual stimulation* includes all potentially reading-relevant stimulation which occurs outside the classroom. Examples include "television," "street signs," and "evening reading hour." When teachers made reference to stimulation (e.g., quality literature), but were not clear whether such stimulation occurred inside or outside of the classroom, the coder examined adjacent nodes in the map for cues as to the proper coding category.

In the few instances where no such cues were available, the coder included the node in both categories. Fourth, *concrete examples of instructional strategy* includes specific things teachers do in the classroom to teach reading. Examples include "read to class" and "group discussion of books". Fifth, *instructional outcomes* included goals of instruction, including "letter-sound correspondence" and "left-to-right scanning". Ratio scores were generated by dividing the number of nodes in a given category by the number of nodes in that map.

Procedure

Participants completed the TORP in September 1990. Following this, they generated concept maps in September 1990, December 1990, and April 1991. As in Beyerbach (1988), subjects generated both the content and structure of their own content maps. This method contrasts with those studies in which the experimenter provides terms for categorization (e.g., Strahan, 1989) or provides a list of supplementing terms spontaneously generated by subjects (e.g., Roehler et al., 1990). Concept maps were scored by the second author in close consultation with the first author. A second administration of the TORP took place in May 1991.

Results

Interrater Reliability

Interrater reliabilities were calculated for each of the 10 dependent variables in the current study. A second coder scored 14 cognitive maps after brief training by the initial coder, and her scores were correlated with those of the initial coder. Interrater reliabilities for each measure are shown in Table 1. As can be seen, interrate reliabilities were acceptably high for all measures in the current study.

Overview

Table 2 presents means, sample sizes, and standard deviations, at time 1, time 2, and time 3, for all variables examined in the current study. Note that the data in Table 2 refer to *all* subjects in the current study, including those who were not present at all three time-periods. Therefore, these data are not useful for examining change across time-periods, as different individuals may be represented at time 1, time 2, and time 3.

Table 1

Interrater Reliabilities for all Dependent Measures

	Correlation	*p*
Measure of structure		
Number of nodes	> .99	<.0001
Number of links	.99	<.0001
Number of chunks	.96	<.0001
Average depth	.95	<.0001
Total depth	.99	<.0001
Measures of content		
Affective state	.98	<.0001
Individual differences	.97	<.0001
Outside class Stimulation	.99	<.0001
Instructional strategy	.95	<.0001
Instructional outcomes	.99	<.0001

N = 14.

In the sections which follow, we first present the results of the correlational analyses, as these are most directly relevant to the primary hypotheses of the current study. We then present the results of both within-subject and between-subject Analyses of Variance (ANOVAs).

TORP Scores

As in previous studies using the TORP, subjects' phonics scores correlated highly with their skills scores, $r(32) = .84$, $p < .0001$ and inversely with their whole-language scores, $r(32) = -.65$, $p = .0001$. For the analyses in the current study, a TORP summary score was generated by subtracting phonics scores from whole-language scores, for each subject. Summary scores correlated highly with subjects' whole language scores, $r(32) = .89$, $p < .0001$, and inversely with their phonics scores, $r(32) = -.93$, $p < .0001$.

Administration of the TORP a second time, at the completion of the study, confirmed that the TORP has a high test-retest reliability, $r(25) = .81$, $p < .0001$, for summary score.

TORP Scores and Teachers' Concept Maps

Structural measures. We predicted no relationships between theoretical orientation and measures of structure. Examination of correlations between TORP summary scores and five

Table 2

Means, Sample Sizes and Standard ***Deviations* for *all*** Dependent ***Measures in the Current Study***

Structural measures	*n*	*M*	*SD*
Nodes			
Time 1	23	28.22	9.45
Time 2	31	30.10	12.16
Time 3	33	26.88	16.33
Links			
Time 1	23	31.91	8.86
Time 2	31	32.17	11.79
Time 3	33	28.42	16.10
Chunks			
Time 1	23	4.74	2.49
Time 2	31	5.32	2.46
Time 3	33	3.97	2.71
Average depth			
Time 1	23	1.48	.43
Time 2	31	1.59	.51
Time 3	33	1.33	.36
Total depth			
Time 1	23	58.52	27.31
Time 2	31	66.13	30.28
Time 3	33	52.82	33.49
Content measures	***n***	***M (%)***	***SD***
Affect			
Time 1	23	10.5	9.7
Time 2	31	8.6	8.9
Time 3	33	7.8	1.9
Individual differences			
Time 1	23	3.0	4.0
Time 2	31	4.2	5.7
Time 3	33	5.0	9.2
Outside class			
Time 1	23	22.1	19.7
Time 2	31	21.9	18.4
Time 3	33	25.5	15.3
Instructional strategy			
Time 1	23	23.4	16.7
Time 2	31	24.1	15.6
Time 3	33	27.3	16.0
Outcomes			
Time 1	23	13.9	12.7
Time 2	31	10.4	8.5
Time 3	33	10.5	11.9

Note: Percentages across content measures will not sum to 100% for a given time period because some nodes were not coded into one of our five content categories.

measures of structure yielded no significant correlations at time 1, 2, or 3. The strongest single correlation (of 30 possible correlations) was between time 1 TORP and number of links at time 1, $r(23) = .35$, $p = .10$ (two-tailed test). No other TORP/structure correlations approached significance. These results suggest that theoretical orientation toward reading predicts little about the structure of teachers' cognitive maps.

Content measures. We predicted that teachers with whole language theoretical orientations would emphasize affect, individual differences, and outside class activities more than teachers with phonics orientations. Support for these hypotheses would come from positive correlations between TORP summary scores and category scores. As can be seen in Table 3, for both affect and individual differences, all six correlations were in the predicted direction, two reached statistical significance (.05 level, two-tailed tests), and two approached significance. These correlations provide moderate support for the hypotheses for affect and individual differences, but not for outside class activities. In fact, correlations for outside class activities were in the opposite direction from predictions, significantly so at time 1. No consistent relationships emerged between TORP scores and other content measures.

Stability of Teachers' Maps

Structural measures. We predicted that the structure of teachers' concept maps would be stable across time. This would yield significant positive correlations within structural measures between times 1, 2, and 3.

As can be seen in Table 4, number of nodes, total depth, number of chunks, and number of links were quite stable across time, with 7 of 12 correlations reaching statistical significance (.05 level, two-tailed tests) and four others approaching significance. Average depth proved stable only between time 2 and time 3. Given the relatively small N values in each cell, these results taken together provide strong support for the stability of structural aspects of concept maps over time.

Content measures. As with structural measures, we predicted that content measures would be relatively stable across time. This would yield significant positive correlations within content category scores between times 1, 2, and 3.

As can be seen in Table 5, all correlations but one were in the predicted direction. However,

Table 3

Correlations Between Two TORP Summary Scores and Three Content Measures at Three Times

Measure		Time 1	Time 2	Time 3
Affect	T1	.33	.26	.22
	TORP	(n = 23, p = .12)	(n = 21, p = .26)	(n = 19, p = .37)
	T2	.63	.44	.34
	TORP	(n = 18, p = .01)	(n = 24, p = .03)	(n = 22, p = .13)
Individual differences	T1	.41	.I4	.34
	TORP	(n = 23, p = .05)	(n = 21, p = .56)	(n = 19, p = .15)
	T2	.50	.34	.23
	TORP	(n = 18, p = .03)	(n = 24, p = .ll)	(n = 22, p = .29)
Outside class	T1	– .48	-.01	-.07
	TORP	(n = 23, p = .02)	(n = 21, p = .97)	(n = 19, p = .78)
	T2	– .46	.04	.15
	TORP	(n = 18, p = .05)	(n = 24, p = .84)	(n = 22, p = .50)

only two reached statistical significance (.05 level, two-tailed tests): time 2 with time 3 for affect, and time 1 with time 2 for individual differences. The time l/time 3 correlations also approached significance for both affect and outcomes. Taken together, these results provide some support for the stability of affect and, to a lesser degree, individual differences as content measures in teachers' concept maps.

Teacher Experience and Teachers' Mups

Structural measures. We predicted that more experienced teachers would generate deeper and better differentiated maps. The data failed to support this hypothesis. However, there were trends toward *inverse* correlations between years of teaching experience and number of chunks at time 1, r(19) = -.38, p = .ll, time *2, r(20)* = –.34, p = *.13,* and time 3, r(19) = -.35, p = .15 (two-tailed tests). No other correlations between years of teaching experience and structural measures approached significance.

Content measures. We predicted that more experienced teachers would generate a higher percentage of nodes related to instructional strategy, affective factors, and individual differences. The data failed to support these predictions. In fact, years of teaching experience correlated *inversely* with affect at time 3, r(17) = – .48, p = *.05,* and, as a trend, at time 2, r(21) = -.34, p = .14 (two-tailed tests). Further, years of teaching correlated inversely with individual differences at time 2, r(21) = -.40, p = .07 (two-tailed tests). No other correlations between years of

Table 4

Correlations Within Five Structural Categories at Three Times

Measure	Time l/Time 2	Time l/Time 3	Time Z/Time 3
Number of nodes	.I3 (n = 20, p = .0003)	*.49 (n = 16, p = .05)*	.39 (n = 21, p = .08)
Links	.72 (n = 20, p = .0004)	*.37 (n = 16, p = .16)*	.37 (n = 21, p = .10)
Chunks	.56 *(n = 20, p = .01)*	*.38 (n = 16, p = .15)*	.62 (n = 21, p = .003)
Average depth	-.12 (n = 20, p = .62)	-.11 *(n = 16, p = .68)*	.57 (n = 21, p = .0006)
Total depth	.58 (n = 20. p = .007)	*.30 (n = 16, p = .25)*	.51 (n = 21, p = .02)

Table 5

Correlations Wiltin Five Content Categories ***at*** Three Times

Measure	Time l/Time 2	Time l/Time 3	Time Z/Time 3
Affect	.21 (n = 20, p = .37)	.42 (n = 15, p = .12)	.70 (n = 21, p = .0005)
Individual differences	.49 (n = 20, p = .02)	.25 (n = 15, p = .38)	.01 (n = 21, p = .96)
Outside class	.30 (n = 20, p = .28)	.14 (n = 15, p = .61)	-.18 (n = 21, p = .42)
Instructional strategy	.24 (n = 20, p = .31)	.29 (n = 15, p = .29)	.13 (n = 21, p = .57)
Outcomes	.25 (n = 20, p = .28)	.40 (n = 15, p = .14)	.14 (n = 21, p = .54)

experience and content measures approached significance.

Effects of Time and Grade

Longitudinal change. Repeated measures ANOVAs were run on all content and structural measures to look for changes from time 1 to time 3. These ANOVAs allow us to examine whether teachers, as a group, changed over time on these measures. The stability correlations, reported previously, allow us to examine whether teachers are consistent, relative to each other, in their scores on these measures.

Table 6

Means for Four Structural Measures at Three Times

Measure	*n*	*M*	*SD*
Nodes			
Time 1	14	30.71	9.79
Time 2	14	32.64	11.71
Time 3	14	24.93	10.82
Links			
Time 1	14	34.36	8.33
Time 2	14	36.64	10.73
Time 3	14	27.71	10.13
Chunks			
Time 1	14	5.36	2.24
Time 2	14	5.79	3.22
Time 3	14	4.36	3.10
Total depth			
Time 1	14	67.36	29.36
Time 2	14	75.29	35.76
Time 3	14	51.71	25.20

As can be seen in Table *6*, there was a similar pattern for four of five structural measures: a slight rise between time 1 and time 2 followed by a sharper drop between time 2 and time 3. These patterns were significant for number of nodes, $F(2,26) = 4.60$, $p = .02$; number of links, $F(2,26) = 6.95$, $p = .004$; and total depth, $F(2,26) = 4.02$, $p = .03$; and approached significance for number of chunks, $F(2,26) = 2.25$, $p = .13$. There was no such effect for average depth, nor for any of the content measures.

We also ran a repeated measures ANOVA to see if teachers' TORP scores changed over the course of the study. Results revealed a nonsignificant ($p = .09$) trend toward whole language shift across the time of the study. No other repeated measures ANOVAs yielded results approaching significance.

Effects of grade. Between subjects ANOVAs were run to look for possible effects of grade (kindergarten, first-, and second-) and school on all dependent measures. No statistically significant results emerged from these analyses.

Discussion

The primary aim of the current study was to see if concept maps would reveal cognitive differences between teachers whose theoretical orientations toward the teaching of reading were known to differ. We predicted that teachers known to emphasize whole language would gen-

erate a higher proportion of nodes demonstrating concern for individual differences between children, affect toward the learning of reading, and outside class activities associated with reading. This hypothesis was supported for affect and individual differences, but not for outside class intellectual stimulation. In fact, at time 1, phonics-oriented teachers generated maps with a greater proportion of extracurricular references than did whole language-oriented teachers.

These results suggest that concept maps are useful as a measure of theoretical orientation, and they appear to validate at least two of the assumptions concerning instructional practices which are implied by whole language versus phonics orientations.

The negative findings for outside class intellectual stimulation are somewhat difficult to explain. Examination of individual maps reveals that most teachers generated nodes in this category, including "television," "street signs," "menus," and "parents read to children," just to mention a few. It may be that the importance of extracurricular stimulation, and especially parent input, is so widely acknowledged (Taylor, 1983), that theoretical orientation does not generally affect its salience to teachers. (This does not, however, explain the inverse findings at time 1.)

Similarly, concrete examples of instructional strategy and instructional outcomes are of importance to all teachers, and this may account for the lack of a relationship between these variables and theoretical orientation. The content categories chosen for the current study were exploratory, and we suggest that future researchers cast aside those which fail to bear fruit as they add new ones relevant to their hypotheses. We suspect that crafting narrower categories that more closely match the phenomena under study would increase the probability of finding significant relationships.

As predicted, there were no structural differences between the maps of whole language- and phonics-oriented teachers. Of 30 possible correlations (2 TORP administrations x 3 concept map administrations x 5 measures), only one approached significance, with whole language teachers tending to generate slightly more links than phonics teachers. However, given that we would expect 3 of 30 correlations to be significant at the .10 level by chance, we attribute little relevance to this finding.

Our second hypothesis was that both content and structure of teachers' maps would be relatively stable over time. For content, there was mild support for this hypothesis, with 2 of 15 possible correlations reaching significance (although all but one were in the predicted direction). However, it is more encouraging that the most reliable two content categories, affect and individual differences, were the same two which differentiated whole language and phonics teachers.

For structure, number of nodes, total depth, number of chunks, and number of links all proved relatively stable, while average depth did not. As mentioned previously, average depth was exaggerated in maps in which teachers used few level one and level two nodes, then branching out into a more tree-like structure. The removal or addition of one low-level node (e.g., a solitary level two node which links level one and level three nodes) can radically alter a map's score on this measure. Total depth was generated in response to this problem, though total depth correlates highly with number of nodes. We suggest caution in the use of average depth in future research using concept maps.

Finally, we sought to examine the relationship between teacher experience and concept maps. Based on findings of Beyerbach (1988), Roehler et al. (1990), and Strahan (1989), we predicted that more experienced teachers would generate better-differentiated maps with greater attention to child affect, individual differences, and instructional strategies.

Surprisingly, we found inverse correlations between years of teaching experience and affect at time 3 and, as a trend, at time 2. Similarly, we found an inverse correlation between years of teaching experience and individual differences at time 2. For structural measures, we found trends toward inverse correlations between years of teaching experience and number of chunks at times 1, 2, and 3.

Although we do not have data concerning teacher effectiveness in this study, our results do seem to be consistent with the contention (e.g., Berliner, 1986; Felder, Hollis, Piper, & Houston, 1979) that years of teaching is not an adequate predictor of teacher effectiveness. Barnes (1987) suggests that a curvilinear, but possibly negative overall, relationship might exist between years of experience and teacher effectiveness. To see if a

similar relationship might exist between years of experience and quality of concept maps in the current study, we conducted a post hoc analysis in which we divided years of teaching experience into quartiles and used these quartiles as grouping variables for examination of content and structural measures of the maps.

For content measures teachers with the most experience generally produced maps with the lowest percentage of references to affect, individual differences, and instructional outcomes. However, 1 x 4 ANOVAs revealed these effects to be statistically significant only for individual differences at time 3 and instructional outcomes at time 1.

For structural measures, the findings were more robust. For number of nodes at time 1 and 2, number of links at times 1 and 2, number of chunks at time 2, average depth at time 1, and total depth at times 1 and 2, teachers with the most experience generated maps with the lowest scores. The 1 x 4 ANOVAs revealed these effects all to be significant at the .05 level. Although these results suggest a sort of "burnout" occurring with years of experience, our results do not allow us to discriminate between burnout toward teaching and burnout toward participating in research.

Burnout toward the research process is clearly implicated by the results of the repeated measures ANOVAs, which showed that teachers' maps in general became larger and more elaborate between times 1 and 2, then declining severely on all structural measures between times 2 and 3. Two teachers at time 3 wrote notes saying their maps would be abbreviated and two refused to generate maps at this time (all four of these teachers' time 3 data were dropped from the analyses). Such burnout may also explain the finding (see Table 3) that the most reliable correlations between TORP summary scores and content measures occurred when time 1 maps were used. It appears that it is stretching things to ask teachers to generate more than two concept maps in a l-year period (and that their initial map may be, in some sense, their most "accurate"). Although the structural aspects of teachers' maps declined significantly between times 2 and 3, there were no significant changes in content measures (recall that content measures are percentage scores).

The nonsignificant change toward whole language orientation across the course of the study may reflect the effects of the inservice training teachers were receiving during that time. We doubt, however, that teachers' cognitive maps were strongly influenced by the inservice program. One of the workshops concerned "Cultural Diversity in the Classrooms," and might have been thought to have an influence on the emphasis teachers gave to individual and developmental differences on the final maps. This was not the case, although the absence of a control group renders speculative any discussion of the effects of inservice training.

In conclusion, we have shown that concept maps, or ordered trees, are useful for discriminating between teachers who hold different assumptions about the teaching of literacy. Further, we present evidence suggesting that content, as well as structure, of teachers' concept maps should be examined and that it is useful to allow teachers to spontaneously generate their own nodes. Our results suggest that affect and individual differences are two dimensions along which teachers differ as a function of theoretical orientation — in future research, it would be helpful to uncover other dimensions which reliably discriminate whole language- and phonics-oriented teachers.

In a separate research project (Mergendoller, Sacks, Clement-Glass, & Horan, 1994), we are currently examining the relationship between proclaimed theoretical orientation, classroom practices, and student outcomes. We will be able, upon completion of this study, to assess the extent to which cognitive maps of our teachers accurately reflect their practices in the field, and the effects of those practices on their young students.

Note

[1] Teacher concept maps were completed at regularly scheduled meetings. Unlike studies conducted in teacher preparation programs, teachers were not required to attend the meetings, nor were they required to participate in the current study. Not all teachers were present at all meetings, and some teachers who were present nonetheless chose not to participate in the study on certain occasions. This explains the fact that for only 14 teachers were maps available on all three occasions. For all analyses which did not require comparisons across time, all available concept maps were used, maximizing the power of statistical tests while minimizing the possibility of a selection bias.

References

Barnes, J. (1987). Teaching experience. In M. J. Dunkin **(Ed.),** ***The international encyclopedia of teaching and teacher education*** (pp. 6088612). Oxford: Pergamon Press.

Berliner, D. C. (1986). In pursuit of the expert pedagogue. ***Educational*** Researcher, 15(7), 5-13.

Beyerbach, B. A. (1988). Developing a technical vocabulary on teacher planner. Preservice teachers' concept maps. ***Teaching and Teacher Education. 4. 3399347.***

Chall, J. (1967). ***Learning to read: The great debate. New*** York: McGraw-Hill.

Clark, C. M., & Peterson, P. L. (1986). Teachers' thought processes. In M. C. Wittrock (Ed.), ***Handbook of research on teaching*** (pp. 2555296). New York: Macmillan.

Deford, D. E. (1985). Validating the construct of theoretical orientation in reading instruction. ***Reading Research Quarterly, 20, 351-367.***

Dewey, J. (1929). Sources of ***a science*** **of** *education.* ***New*** York: Liveright.

Felder, B. D., Hollis, L. Y., Piper, M. K., & Houston, W. R. (1979). ***Problems and perspectives of beginning teachers: A follow up study.*** University of Houston, Houston, TX.

Goodman, K. (1986). ***What's whole in whole language?*** Portsmouth, NH: Heinemann.

Heald-Taylor, G. (1989). ***The administrator 's guide to whole language.*** Katonah, NY: Richard C. Owen Publishers.

Mergendoller, J. R., Sacks, C. H., Clement-Glass, E., & Horan, C. (1994). ***The development of kindergarten reading skills: A study of the relationship of teachers' beliefs about reading instruction. classroom process, and student outcomes.*** Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Naveh-Benjamin, M., McKeachie, W., Lin, Y., & Tucker, D. (1986). Inferring students' cognitive structures and their development using the "ordered tree technique". ***Journal of Educational Psychology, 78,*** 130-140.

Piaget, J. (1952). ***The origins of intelligence in children.*** International Universities Press.

Roehler, L. R., Duffy, G. G., Conley, M., Herrmann, B. A., Johnson, J., & Michelson, S. (1990). ***Teachers ' knowledge structures: Documenting their development and their relationship to instruction.*** Unpublished paper, Institute for Research on Teaching, Michigan State University.

Smith, F. (1988). ***Joining the literacy club: Further essays into education,*** Portsmouth, NH: Heinemann.

Strahan, D. B. (1989). How experienced and novice teachers frame their views of instruction: An analysis of semantic ordered trees. ***Teaching and Teacher Education. 5, 5347.***

Taylor, D. (1983). ***Family literacy: Young children learning to read and write.*** Portsmouth, NH: Heinemann Educational Books.

Submitted 23 February 1993
Accepted 17 February 1994