ELSEVIER

Poetics 22 (1994) 291–312

POETICS

Extracting culture through textual analysis *

Kathleen Carley

Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Abstract

Language has been viewed as a window on the mind. Language is also a window on culture. Through analyzing texts the interplay between human cognition and culture can be examined. Through analyzing texts cognitive similarities and differences across individuals, which serve as a basis for culture can be described. Through analyzing texts the impact of culture on individual behavior can be examined. In addition, such analyses can locate similarities and differences across cultures and changes within cultures. This paper explores the relative benefits for using content analysis and map analysis for extracting and analyzing culture given a series of texts. Content analysis has been the traditional textual analysis method used for examining culture. However, it is not theoretically grounded. In contrast, map analysis has received less use and is theoretically grounded in an understanding of human cognition. It is shown that under certain conditions map analysis subsumes content analysis. Researchers can thus use map analysis not only to extract and analyze culture but to examine the relationship between cognition and culture. Illustrative applications are drawn from four different studies.

1. Introduction

Language is frequently viewed as a window on the mind, particularly among researchers interested in human cognition and problem solving. Various textual analysis techniques have been proposed for locating information on human cognition given a text; e.g., linguistic content analysis (Roberts, 1989), semantic grammars (Franzosi, 1990a,b), mental model analysis (Carley and Palmquist, 1992; Carley, 1993), protocol analysis (Ericsson and Simon, 1984), procedural task analysis (VanLehn and Garlick, 1987). Within this research tradition, concern with knowledge representation has led researchers interested in examining mental models to turn to textual analysis techniques that examine not only the concepts in

* Thanks go to Aaron Cicourel for providing the children's recall data used in this study and to Tom Dale for doing the initial coding of the robot data.

SSDI 0304-422X(93)E0012-G

the texts but the inter-relationships among them. Researchers who use these techniques often focus on extracting the text's author's mental model of some construct.

Language, however, is also frequently viewed as a window on culture. Indeed natural languages are fraught with terms whose meaning is embedded in the culture and which reflect the extant culture. In the social and political sciences where the interest in culture is perhaps most pronounced, the most common type of textual analysis used to examine culture has been, one form or another of content analysis (North, 1963; Gerbner, 1969; Holsti, 1969; Gottschalk and Gleser 1969; Gottschalk and Winget, 1969; Janowitz, 1969; Stone et al., 1968; Berelson 1952; Fan, 1988; Namenwirth and Weber, 1987; Garson, 1985; Neuman, 1989; Weber, 1984). Other related textual analysis techniques that have also been used include proximity analysis (Danowski, 1980, 1982, 1988; vanMeter and Mounier 1989), and concordance analysis (Young et al., 1936). Within this research tradition, the focus has been on concepts and their distribution within and across texts. Researchers who use these techniques often focus on extracting word counts associated with some construct.

While content analysis has been fruitfully applied to understanding culture in the large (as a composite across individuals) there has been little success in using content analysis to understand individual cognition and the role of culture at the individual level (Weber, 1984; see also Roberts, 1989). Textual analysis techniques based on the inter-relationships among concepts have been used successfully to examine individual cognition. In this paper, it is demonstrated that textual analysis techniques that consider the distribution across and within texts of concepts and the inter-relationships among concepts can be fruitfully used to examine the role that culture plays in human memory and cognition.

In this paper, I examine the relative benefits of two textual analysis techniques for analyzing the interplay between culture and cognition in texts. The two techniques are traditional content analysis (focusing simply on concepts and their frequency) and map analysis (focusing on concepts, the inter-relationships among them, and the frequency of concepts and inter-relationships). See Carley (1993) for an overview and discussion of the differences between content analysis and map analysis. These two techniques are demonstrated through a re-analysis of data drawn from four studies: the role of robots in science fiction (Thomas, 1991), the difference between drama and comedy in modern American culture (Carley and Kaufer, 1993), the ability of children to recall a story (Cicourel and Carley, 1990; Cicourel et al., 1992); and the evolution of student's decision maps (Carley, 1984, 1986). This study shows that by attending to both concepts and the inter-relationships among them when analyzing texts a better understanding of the relationship between individual cognition and culture can be achieved.

There are two theoretical premises underlying this study. [1] First, texts contain

[1] See Carley and Palmquist (1992) for a more detailed discussion of the theoretical basis of map analytic techniques.

an encapsulation of a portion of the author's mental model at the time the text was composed. Thus textual analysis techniques can be used to extract samples of the author's mental model. This claim, while not completely novel is fairly uncommon in the social sciences where the role of the author is typically taken for granted or ignored (Kaufer and Carley, 1993). Second, individual mental models provide insight into the relationship between individual cognition and culture as both cognition and culture evolve through, and concurrent with, the evolution of language. The second of these claims is certainly not new to the social sciences. Mental models are central to social theories in which individuals represent the world and interact with it through symbols. See for example the work of Mead (1962; 1964) and the symbolic interactionists (Blumer, 1969; Stryker, 1980). Similarly, mental models are central to social theories in which individuals create the social world through their use of symbols. See for example the work of social constructivists (Latour and Woolgar, 1979; Knorr-Cetina, 1981). The purpose here is not to belabor the theoretical underpinnings of the map analytic approach, which links mental models to culture, but to simply point out that were we able to link mental models to culture we would have the requisite technology for empirically examining a wide range of claims made by various social theories.

2. Methodologies

Content analysis is a family of techniques all of which emphasize frequency counts of concepts within and/or across texts. Additional details on content analysis can be found in Weber (1984). When content analysis is used the researcher simply counts whether (or how often) the concepts of interest occur in (or are inferred to be in) each text being analyzed. Regardless of whether such categorization occurs, the result is essentially a concept by text matrix.

Map analysis is a family of techniques all of which emphasize situated concepts and the relationships among them (Carley, 1993). The particular type of map analysis used in this paper is the mental model extraction technique proposed by Carley. Additional details on map analysis can be found in Carley (1988, 1993) and Carley and Palmquist (1992). When map analysis is used the researcher simply counts whether, for each possible pair of concepts of interest, the pair of concepts form a statement that occurs in (or is inferred to be in) each text being analyzed. A pair of concepts form a statement if they are part of a semantic unit such as 'robots are nice', 'comedies can be satires', 'boy walks'. Within a statement there is a 'link' between the two concepts. This link can be directed (e.g., from the first concept to the second), signed (e.g., a positive link such as a is a b or a negative link such as *a* is not a *b*), or possess a certain strength (e.g., a degree of emphasis). When map analysis is used there are essentially two results, a statement by text matrix and a concept by text matrix.

Researchers interested in using either content analysis or map analysis must make a variety of choices that in turn affect the interpretation of the results (Carley, 1993). Thus, it is important that the specific coding choices be specified.

In this paper, the following coding choices are made: When using either content analysis or map analysis the researcher may choose to classify the concepts into a set of categories or 'types' of concepts. Herein, all concepts are treated as being of the same type. For either content analysis or map analysis the researcher may choose to pre-define a set of concepts or let the text dictate the concepts. Herein, for each data set, the collection of concepts used are extracted as the texts are analyzed. For either content analysis or map analysis the researcher may choose to use only concepts that appear explicitly in the text, to infer concepts from the texts, or to use a generalized set of concepts some of which appear in the text and some of which are inferred. Herein, the latter approach of using a set of generalized concepts is used. When content analysis is used the researcher must choose whether to count 'if a concept appeared in a text' or 'how often did the concept appear in the text'. In this paper, I simply count: did the concept appear in (or can it be inferred from) the text. When map analysis is used the researcher may choose to classify the statements into a set of categories or "types" of statements. Herein, all statements are treated as being of the same type. Additionally, when using map analysis the researcher must choose whether or not the statements are directed and whether or not they have a strength. In this paper, all statements can be either uni- or bi-directional and all statements are of the same strength. In other words, I will be counting simply: did the pair of concepts appear as a statement in (or can it be inferred from) the text. As a final point, since part of the goal of this paper is to contrast content analysis and map analysis: in this paper the set of concepts used are the same whether content analysis or map analysis results are presented. Thus, in comparing the content analysis and map analysis results, the 'value added' of map analysis can be judged simply by seeing what more do you learn when you have statement information as well as concepts information.

To further illustrate the two techniques let us consider how they would be used to code the following text using the concepts robot, humanoid, human, metallic, and soulless. The text is: "Voltan is a robot. Although it looks human it has no soul". The result of the content analysis is that four of the five concepts occur in the text: robot, humanoid, human, and soulless. The result of the map analysis is that there are five statements and four concepts. The statements are "robot is humanoid", "robot is soulless", "robot is not human", "human is humanoid", and "human is not soulless". The concepts are robot, humanoid, human, and soulless.

Note, map analysis subsumes content analysis when the researcher focuses simply on the question "is the concept or statement present in the text?" Thus, the value added of map analysis over content analysis is the extra value of knowing when the 'story told' by statements elaborates on or contradicts that told by only the concepts.

Before turning to the data and the results there is one final methodological point that must be considered – the nature of social knowledge or culture. Using these techniques several aspects of culture can be examined. First, culture as shared or social knowledge can be located. Social knowledge was implicitly defined by Polanyi (1958, especially, pp. 216–219, and 264–266) in his discussion of the pursuit of science as "tacit knowledge" resulting from "tacit consensus" which is

itself a by-product of the articulation of "common experience", and more importantly a product of the transitivity of appraisal across a continuous network of critics. This suggests that social knowledge is common but not necessarily uniformly shared. In this paper, social knowledge is that set of concepts (or statements) such that each concept (or statement) occurs in 50% or more of the texts being analyzed. Second, cultural diversity is measured as the number of concepts (or statements) used in the texts. Third, when map analysis is used cultural density can also be measured. Cultural density, or the degree to which the social knowledge that forms the basis of the culture is interconnected, can be measured as the ratio of statements to concepts. In this case, the higher the number the more interconnected the concepts.

3. Data

To illustrate the relative benefits of content analysis and map analysis in relating cognition to culture four different data sets will be used. The first data set is concerned with the difference between comedy and drama in American culture (Carley and Kaufer, 1993). The second data set is concerned with the portrayal of robots in science fiction over time (Thomas, 1991). The third data set is concerned with the role of culture in children's recall of stories (Cicourel et al., 1992). The fourth data set is concerned with the shifts in individual's cognitions and the role of cultural knowledge in the decision-making process (Carley, 1984, 1986). The purpose in using these data sets is that each illustrates a different point about how the methods of content analysis and map analysis can be used to examine the relationship between cognition and culture. I have ordered them from the furthest removed from the cognition of a single individual to the closest. No claim is being made about the sufficiency of, or quality of, these data for addressing the issue for which they were initially collected. The interest here is not in proving any claims about the relationship between cognition and culture, but in demonstrating how these textual analysis methods can be used to empirically address such claims.

The theater data set is drawn from the electronic thesaurus Word Finder on the Macintosh (Carley and Kaufer, 1993). We extracted first for comedy and then for drama all concepts that occurred in the entry for that concept (level one concepts) and then all concepts occurring in the entry for each of those concepts (level two concepts). If two concepts appear in the same entry they are treated as a statement. The result is a set of 310 concepts and two texts (one each for comedy and drama). This data set allows us to examine average cultural differences in American talk about the theater. Here we are not looking at individual cognition per se, but at the common usage of concepts (and so in some sense the average cognition) in the theater domain.

The robots in science fiction data set is based on 30 texts ranging from Shelley's *Frankenstein* and Karel Capek's *Rossum's Universal Robots* to various science fiction robot books in the 1980s such as Piers Anthony's *Robot Adept*. The texts are divided across three time periods pre-1950s, 1950s–1960s, and 1970s–1980s.

The texts are actual books. From these books, only that portion of the text was coded that served to define the term robot relative to the description of the leading robot or robots. If there was more than a single lead robot, as in Isaac Asimov's *Robots and Empire*, the definition of robot was coded from the perspective of each lead robot (e.g., that of Danielle and Giskard). In this case, a single book is treated as multiple texts, one for each lead robot. This resulted in 30 texts divided as follows – six from the pre-1950s, 12 from the 1950s and 1960s, and 12 from the 1970s and 1980s. For this analysis, the coded texts were recoded from Thomas's coding so that highly similar concepts such as positive and very-positive were combined. As a result all texts are coded using a set of 392 concepts. Recoding also involved adding a series of implied links between key defining terms and the concept robot. This was done by passing all coded texts through the post-processor SKI (Carley, 1988) to make explicit, implicit statements.

The story recall data set is based on a subset of the data collected by Aaron V. Cicourel and David E. Rumelhart as part of a larger project and is based on data from 8- and 9-year-olds tested in a combined 2nd and 3rd grade class and in a laboratory by Aaron Cicourel and David Rumelhart. This work is reported on by Cicourel and Carley (1990) amd Cicourel et al. (1992). In this paper, I use only the data drawn from two texts for each of 4 of the 28 eight and nine-year-old children. The two texts are the initial story description and last recall. The initial story description is the discourse the child carries on while reading the story. The last recall is the child's recall of the story 10 to 12 weeks after having read the story. The data reported here is a recoding of the pragmatically coded data reported by Cicourel et al. (1992). In making this recoding all pronouns were converted to nouns. There are 79 concepts in this recoded data set.

The decision-making data set is based on a subset of the data collected by Carley as part of a larger study on the decision-making behavior of a group of undergraduates selecting a new tutor for their living group (Carley, 1984). Carley interviewed the students multiple times throughout the tutor selection process. These interviews were then coded as maps and run through the post-processor SKI to improve the coding (Carley, 1988). In this paper, I use the modified maps for those five students for which there are maps at time 1 (beginning of process), time 2 (middle of process), and time 3 (end of process). There are 217 concepts in this data set.

4. Cognition and culture

4.1. Culture as practiced – The case of American comedy and drama

In America, comedy is a more developed concept than is drama according to this thesaurus (see Table 1). In America, there is more social knowledge/diversity regarding comedy and greater density than there is regarding drama. Note, since the data is drawn from the thesaurus, which can be viewed as a compendium of social knowledge, there is no difference between the measure of social knowledge

Table 1
Portrayal of comedy and drama

	Social knowledge		Cultural density
	Concepts	Statements	
Comedy	254	749	2.95
Drama	151	358	2.37

and cultural diversity. In fact, there are 103 more concepts used to describe comedy than drama and 391 more statements. These differences are both significant using a two-tailed t-test at the 0.05 level: concepts ($t = 163.298$, $df = 618$) and statements ($t = 3649.04$, $df = 191{,}578$). [2]

Moving on to Table 2 we see that not only is there simply more knowledge about comedy, the knowledge about comedy is very different than that about drama. Only one-third of the terms used to discuss drama and less than 50% of the statements are unique to drama. However, for comedy close to two-thirds of the terms and 75% of the statements are unique. The greater uniqueness in statements than in concepts suggests that while it may be more common in America to describe the theater in common terms (use the same concepts), the meaning of those terms is altered as one moves from a discussion of comedy to one of drama. And, moreover, the meaning of those terms is much more elaborate within discussions of comedy.

This suggests, for example, that within the American culture humor is a less constrained idea than is being serious. This can be seen by looking at the concepts

Table 2
Information on comedy and drama

	Concepts	Statements	Cultural density
Comedy only	154	558	3.62
Drama only	51	167	3.27
Comedy and drama	100	191	1.91

[2] The test based on concepts assumes that the presence or absence of a concept is a binary variable. Concepts are assumed to be distributed uniformly within a text with a probably of P/N where P equals the number of concepts that occurred in the text and N equals the number of concepts possible (e.g., 310). The number of concepts present in a text is assumed to be normally distributed. The t-test is simply $P1 - P2$ $((P1\ (1 - P1)) + (P2\ (1 - P2))1/2$. For statements a similar procedure is followed. In this case, however, N is the number of possible statements which is calcualted as the number of concepts possible times the number of concepts possible minus 1. This assumes that it is possible to create a statement with each pair of concepts and that direction matters. This test assumes that the presence of one statement is independent of the presence of another statement, an assumption which may or may not be warranted.

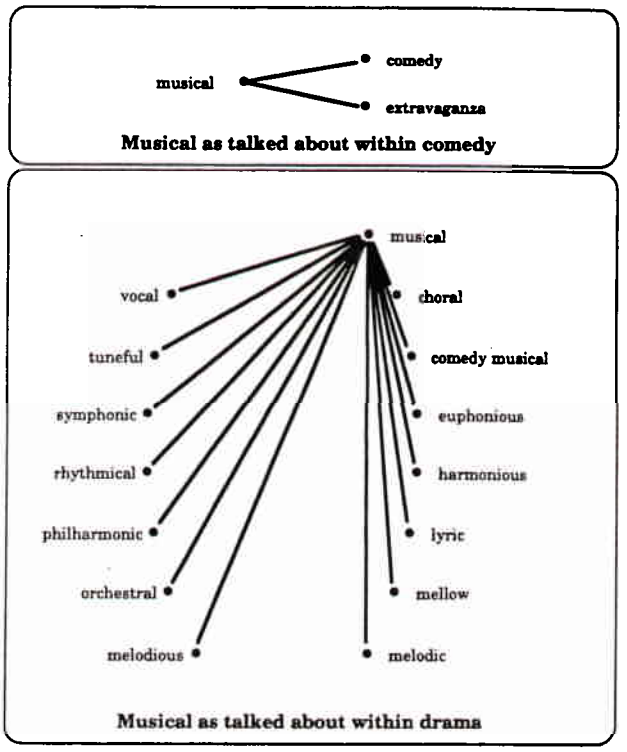


Fig. 1. Musical as defined from the comedy and drama perspective.

unique to comedy, which include escapade, disrespect, chatter, and wink, whereas the drama list includes adversity, disaster, misfortune, thriller, and puzzle. The difference here may go well beyond culture. Studies of affective terms may simply reveal a wider range of positive terms than negative terms. This would be a point for further study.

It is interesting also to look at the commonalities across the two types of theater. Let us take a content analysis approach. We find that concepts common to both comedy and drama are typically types of theater – farce, duration, epoch, musical, mystery – or terms of art for the field – scene, score, season. By looking at similarities and differences across the concepts in an organizational field such as the theater we can begin to get a feeling for the rhetoric of the organizational field and then we could relate it to the structure of the field. This is one way in which studies of culture could begin to move back into organizational studies.

Moving beyond content analysis to map analysis we note that the definitions of comedy and drama contain subdefinitions of various terms. As previously noted, while talk of comedy and talk of drama may use some of the same terms the meaning of those terms differs depending on the domain. An example of this is the term musical. In Fig. 1, the sub-maps for the term musical within the comedy and

Table 3
Portrayal of robots in science fiction

	Social knowledge		Cultural diversity		Cultural density	
	Concepts	Statements	Concepts	Statements	Social	Diverse
pre 1950s	33	55	162	249	1.67	1.54
1950s and 1960s	26	45	205	321	1.73	1.57
1970s and 1980s	30	54	233	358	1.80	1.54

the drama domain are shown. Within comedy, a musical is a comedy and an extravaganza. In contrast, relative to drama, a musical can be a choral or comedy musical. In addition, and only from the perspective of drama, a musical is euphonious, harmonious, lyric, mellow, melodic, melodious, orchestral, philharmonic, rhythmical, symphonic, tuneful, and vocal. Map analysis thus enables us to move beyond seeing whether the terms of art that define the rhetoric of the organizational field are similar to seeing whether the way these terms of art are used, their meaning, are similar.

4.2. Cultural shifts – The case of robots in science fiction

Science fiction is a genre of literature that has the capability of affecting the public's understanding of, and image of, various scientific artifacts such as robots. The writers of science fiction can be viewed as trendsetters affecting and reflecting cultural perceptions and concerns with science. By examining how an artifact is portrayed by a science fiction writer we gain insight into the author's vision. By examining commonalities in how various science fiction writers portray the same artifact we gain insight into the cultural perception of that artifact.

Over time, the portrayal of robots has become more complex – see Table 3. Over time, the level of social knowledge appears to heave decreased and then returned almost to its original level in terms of concepts and statements. Further, over time the diversity of the culture has increased, both in terms of concepts and statements. In other words, both content analysis and map analysis let us know that the shared vision of what it means to be a robot (social knowledge) has vacillated and the possibilities of what a robot might be (cultural diversity) has become elaborated. This suggests that over time there has been an increase in the complexity with which robots are described, but the shared vision about what is a robot has stayed comparable or even decreased in size. All differences in Table 3 are significant using a two-tailed t-test at the 0.05 level. In addition, map analysis enables us to find that the cultural density in shared knowledge increased over time, but the density in diversity has not increased. In other words, over time the number of concepts used by the majority to discuss concepts has decreased while the degree of interconnection among those concepts has increased. These authors may be using fewer similar words to describe robots but they are using those words in a more similar fashion. With respect to diversity we saw that the number of concepts and statements rose over time, whereas, the level of density remained

Table 4
Unique information on robots

	Concepts	Statements	Density
pre 1950s	66	94	1.42
1950s and 1960s	79	118	1.49
1970s and 1980s	101	144	1.43

fairly constant. This suggests that authors are increasingly making statements about what new features (concepts) robots have.

In Table 4 information on concept and statement usage that is unique to that time period is presented. Over time, the amount of unique information, information that occurs only in that time period (either concepts or statements) is increasing. However, the density is remaining fairly constant. This suggests that in terms of unique knowledge, the authors over time are coming up with a greater variety of images associated with robots. However, these images remain equally porous. Future work might consider whether, in general, there is an average level of density associated with describing any artifact in the science fiction genre, and whether such densities vary across different genres.

Now let us consider the nature of the social knowledge. There is substantial similarity in the socially shared model of robots across all three time periods. In fact 34 statements appear in the social knowledge of all three time periods. In Fig. 2, a map based on the social knowledge that is present in all three time periods is shown. As can be seen, in all three time periods, most authors described robots as being able to walk, talk, hear, reason and think. Robots typically have a face and eyes, and memory. In general there is some agreement that something about robots is good (although there is no agreement about what it is). In addition, robots are of a particular type, take action, have features and emotions, and evoke feelings in the other characters in the story.

Now let us consider what additional knowledge is shared in the pre-1950s. In Fig. 3 the map of the social knowledge that occurs in the pre-1950s but not in both of the latter periods is shown. In Fig. 3 (and 4 and 5) some statements from the map in Fig. 2 have been added as dashed lines. These statements have been included to make the maps more readable. However, the reader should keep in mind that if the line is dashed then that statement was also made in both of the other time periods. Prior to 1950 robots were also described as being non-metallic humanoids able to eat, fight and murder. The other characters in the book treated the robot with anger, disdain and fear, and thought of them as formidable and intelligent. Robots could display emotions such as fear, anger and pain. For these reasons, and others less agreed on by the writers, robots were thought of in a negative light and as being bad. Thus, even though the authors portrayed aspects of the robots as good (Fig. 2), overall, and much more commonly, they were thought of in a negative light.

In Fig. 4 the map of the social knowledge that occurs in the 1950s and 1960s but not in both of the other periods is shown. We see that, in keeping with the

Density
1.42
1.49
1.43

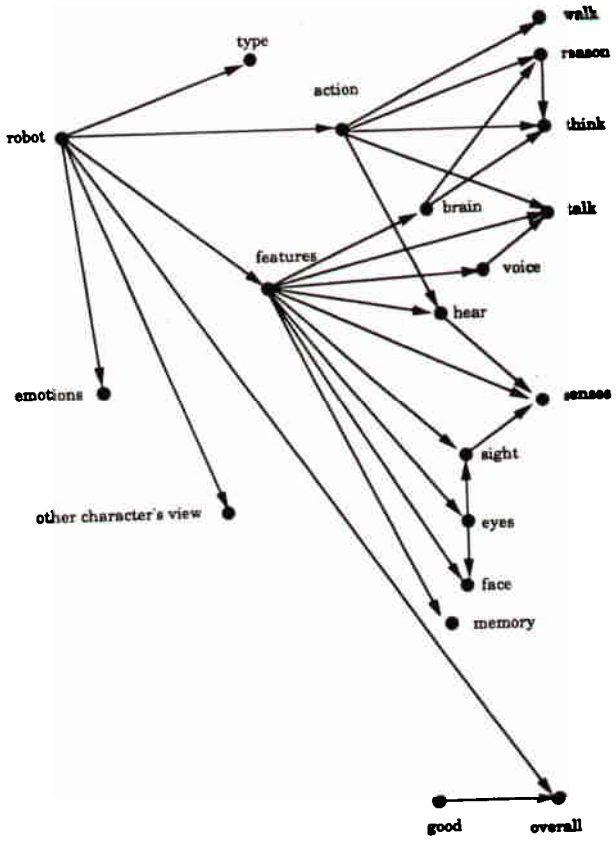


Fig. 2. Definition of robot as shared across three time periods.

pre-1950s, in the 1950s and 1960s, while most writers ascribed some advantages (good) to robots, for the most part they described the robots as being predominantly bad. However, during this period, there is no longer an agreed upon basis for this negative attribution. In addition, in the 1950s and 1960s, the additional characteristics ascribed to robots were less detailed. During this middle period robots came to be viewed as a metallic-humanoid composed of circuits. Other characters in the texts viewed the robots with disdain or were oblivious to them, emotions shared by the robots.

By the 1970s and 1980s the dominant attitude toward robots had reversed. In Fig. 5, a map based on the social knowledge about robots in the 1970s and 1980s is shown. As with the last period, robots are viewed as metallic-humanoids with faces, and as being composed of circuits. Now, however, robots are viewed as having a consciousness and as capable of exhibiting friendship, loyalty, pride and trust. The other characters in the story were able to perceive these emotions in the robots and to treat the robots as friends. Overall, robots during this latter period were viewed in a positive light and as being essentially good.

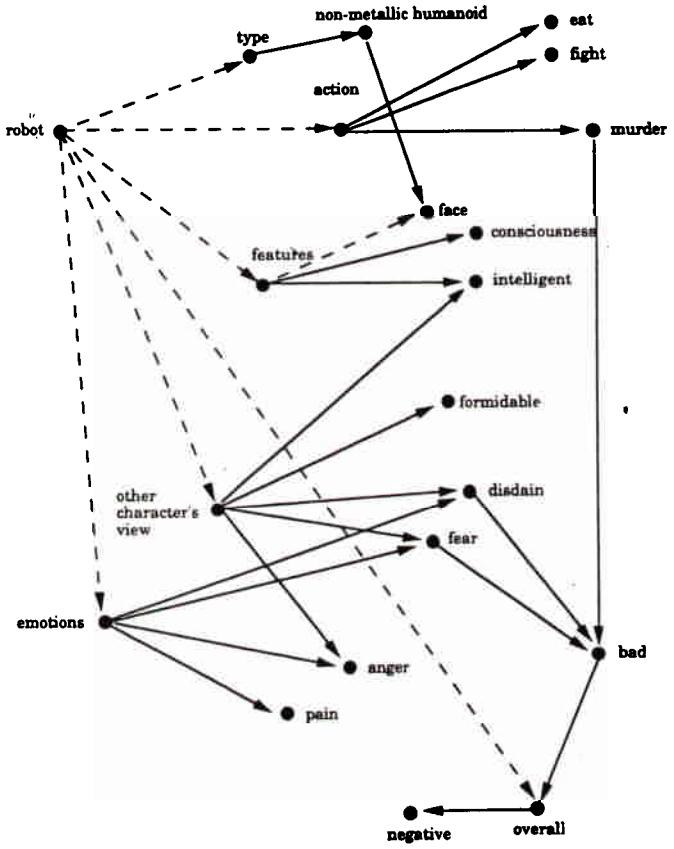
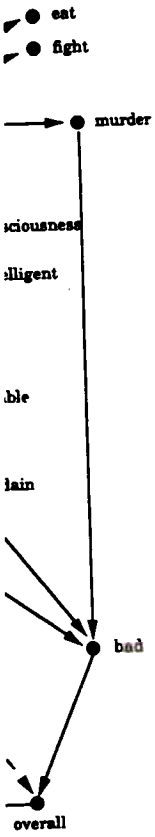


Fig. 3. Shared perception of robot unique to the pre-1950's.

Another way in which to look at the changes in culture is by looking at that information that is unique to the authors of a particular time period. Only in the pre-1950s did characters in the stories view the robots with hatred and pity as demonical-corpses, monsters, soulless and vulnerable. Only in these early years were robots described in terms of batteries, electron tubes, and human parts. Only in these early years were robots described as being able to fly, marry, and reproduce, and as having emotions such as generosity, delight, and despondence.

By the 1950s and 1960s characters in stories started to look upon robots as impressive, infallible, merciless, and superior. Robots were seen as being able to dance and create reports, as exhibiting bravery and self-consciousness, and as having intelligence, being organic, sexless, and having internal power-generators. And only in this middle period were robots described as being able to explode and thereby act as weapons.

The 1970s and 1980s attached yet a different set of unique features to the robots: the ability to analyze emotions, to be charismatic, to contain computer-gen-



e pre-1950's.

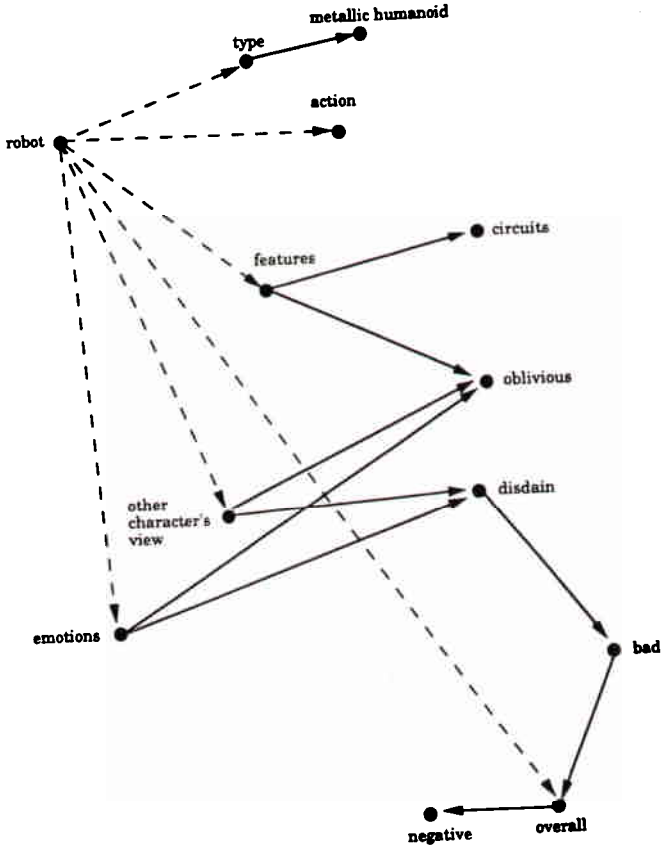


Fig. 4. Shared perception of robot unique to the 1950's and 1960's.

ulture is by looking at that ar time period. Only in the ots with hatred and pity as . Only in these early years ibes, and human parts. Only ig able to fly, marry, and , delight, and despondence. ted to look upon robots as were seen as being able to self-consciousness, and as internal power-generators. as being able to explode and

t of unique features to the ic, to contain computer-gen-

erated cells, and to be masculine. Only in this last period did characters in the books respond to the robots as being alive, clever, loyal, powerful, and obedient. Only during this latter period were robots portrayed with complex emotional reactions such as sarcasm, curiosity, and embarrassment. Only during this period were robots imbued with the ability to make love, program computers, translate texts, and dream.

Clearly, there is less social knowledge about robots across the decades than there is unique knowledge. Nevertheless, sufficient changes pervade the culture that the overall view of robots as portrayed by these authors changed from bad to good. Clearly part of what is unique to each period are technological changes reflecting current science fiction thinking. Another major part of what is unique to each period is the portrayal of how people in the books respond to the robots, the portrayal of the expected culture. Clearly (and to the extent these texts are representative), throughout the 20th century the meaning of robot has changed dramatically from its inspiration, the Czech word robota meaning work or compul-

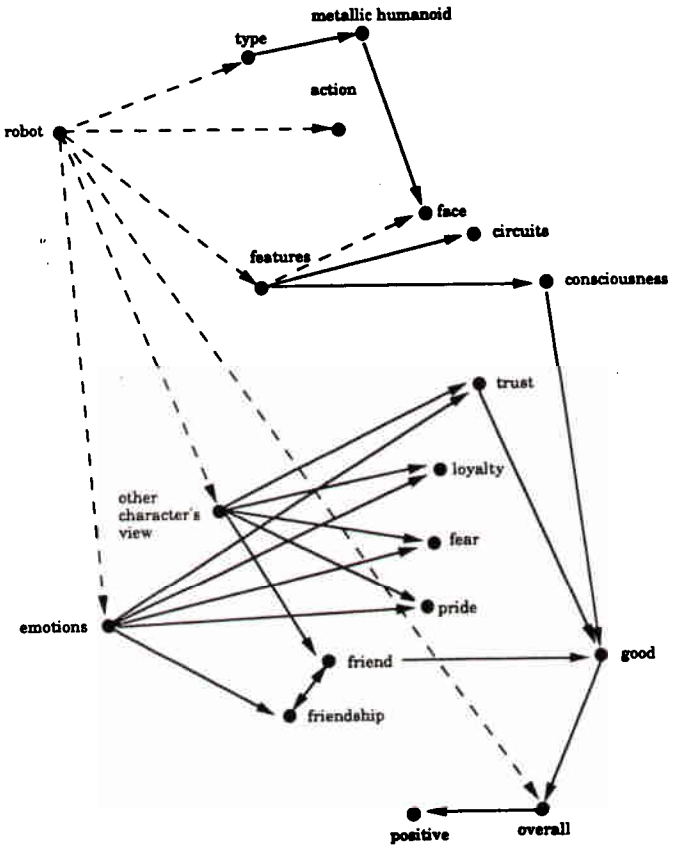


Fig. 5. Shared perception of robot unique to the 1970's and 1980's.

sory service. Future work might consider whether such changes in 'robot' occur in non-fiction work.

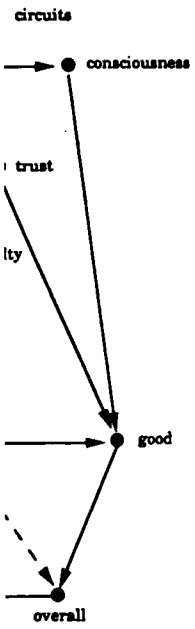
4.3. Culture and recall – The case of children's story recall

The final example is based on a study of a children's recall of stories by Cicourel et al. (1992), and is based on data from 8- and 9-year-olds tested in a combined 2nd and 3rd grade class and in a laboratory by Aaron Cicourel and David Rumelhart. Stories are cultural artifacts. Children are frequently exposed to stories. The ability of the child to recall a story is a function of, among other things, the child's mental models extant in the child's memory. Mandler (1984) suggests that the extent to which people can recall a new story depends on the extent to which the story incorporates regularities in traditional stories. Cultural knowledge about plots and actions increases understanding. Cicourel et al. go on to suggest that children's recall of stories depends on the structure of their cognition and their cultural knowledge.

Table 5
Recall of stories

Concepts	
L	
M	
N	
Y	
Statements	
L	
M	
N	
Y	
Density	
L	
M	
N	
Y	

Cicourel
contain few
forgetting c
connectivity
concepts di
is there mo
In Table
four childr
reading are
all decrease
than the d
structure o
be droppe
the recall
new ones a
were not
(read-aloud
for the por
of stories t
and regen
Indeed,
house and
recalled a
only durin
a sun. Ad



1970's and 1980's.

ch changes in 'robot' occur in

ecall

's recall of stories by Cicourel
ar-olds tested in a combined
Aaron Cicourel and David
are frequently exposed to
a function of, among other
d's memory. Mandler (1984)
a new story depends on the
traditional stories. Cultural
anding. Cicourel et al. go on
ls on the structure of their

Table 5
Recall of stories

	Read-aloud	Recall	Read-aloud only	Recall only	Both
Concepts					
L	41	21	24	4	17
M	35	31	10	6	25
N	43	21	27	5	16
Y	38	26	17	5	21
Statements					
L	56	26	41	11	15
M	52	43	21	12	31
N	64	22	48	6	16
Y	60	34	36	10	24
Density					
L	1.37	1.24	1.71	2.75	0.88
M	1.49	1.39	2.10	2.00	1.24
N	1.49	1.05	1.78	1.20	1.00
Y	1.58	1.31	2.12	2.00	1.14

Cicourel et al. noted that over time children's description of a story tends to contain fewer concepts and statements. This alone suggests a process of either forgetting or increasing difficulty in recall. However, they also noted that the connectivity increases over time. Connectivity is the opposite of density, i.e., concepts divided by statements. Thus, density decreases over time. The question is, is there more going on than simply random forgetting?

In Table 5, a re-analysis of a portion of this data is presented. The protocols of four children as they read the story and their protocols 10 to 12 weeks after the reading are analyzed. As can be seen in Table 5, concepts, statements and density all decrease as noted by Cicourel et al. Over time, the drop in statements is greater than the drop in concepts. This suggests that recall is weakening the overall structure of the story in memory, and is not simply causing a portion of the story to be dropped. However, we also see that the change from the read-aloud period to the recall period involves not only dropping concepts and statements but adding new ones as well. That is, there are concepts and statements used in the recall that were not previously used. Moreover, the density of the forgotten information (read-aloud only) and the generated information (recall only) is much higher than for the portion recalled from one period to the next. This suggests that the portion of stories that are recalled are simple constructs and that details are both forgotten and regenerated.

Indeed, the recalled material includes concepts such as boy, mother, mother's house and statements such as boy has a mother. General cultural knowledge is recalled and present at both the reading and at a later time. Information present only during the reading includes details such as that the country is hot and there is a sun. Additionally, during the reading there are more phrases such as "should

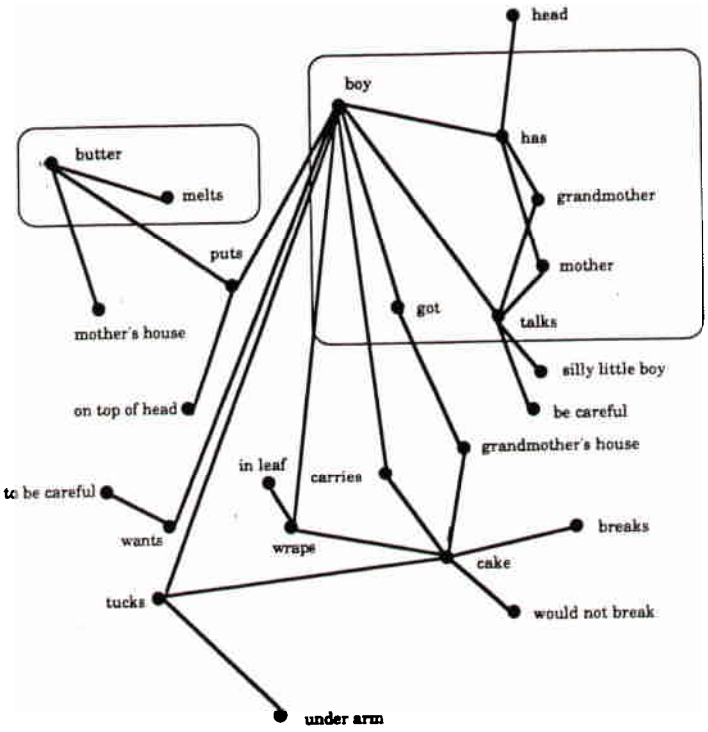


Fig. 6. Shared view of story.

have" and counterfactuals such as "if x then a would not have happened", whereas during the recall elaborations appear such as "the boy did things backwards", "the boy had a hat", "the boy had a coat", or "the boy was dirty".

Now let us consider what information is common to all four students. In Fig. 6 a map of those statements that are shared by all students during the read-aloud stage are shown. It turns out that the statements that are shared by all students during the recall period are a perfect subset of those present during the read-aloud period. Those statements shared during recall are encircled in Fig. 6. Note that over time not only did the number of concepts and statements used by each student diminish, but the number of concepts and statements shared by all students also diminished. Further, we can see that the type of social knowledge that is retained over time is essentially knowledge about the nature of society. That is, all students remember that "butter melts" and that "the boy has a mother" "the boy has a grandmother", "the boy talks", and the "boy has" and "boy got" things. What is common in these children's recall is knowledge about the physical or physiological properties of things and kinship relations. Future work might see whether this behavior is true for all age groups and is independent of the story being read.

2

ndmother's house

Table 6
Change in students' mental models of tutors

	Time 1	Time 2	Time 3
Concepts			
Ian	69	59	32
Lowell	31	55	74
Jaques	33	52	54
Jubal	31	34	32
Hazel	54	57	46
Statements			
Ian	306	208	109
Lowell	103	208	357
Jaques	104	187	230
Jubal	96	102	109
Hazel	220	201	203
Density			
Ian	4.43	3.53	3.41
Lowell	3.32	3.78	4.82
Jaques	3.15	3.60	4.26
Jubal	3.10	3.00	3.41
Hazel	4.07	3.53	4.41

4.4. Culture and decision making – The tutor selection case

The final example is based on a study of students' decision-making behavior by Carley (1984). Individual's decisions are a product of their own view on a situation and their position in the social system. An individual's decision-making process involves two processes: (1) information collection, and (2) evaluation and choice. In this study, the students could collect information over several months before they needed to make a decision. During these months, individuals change their mental models of what they want in the tutor in response to the information they have collected.

In Table 6 changes over time in five students' mental models are shown. As can be seen, students vary dramatically in the number of concepts and statements they use. For each student, density also varies over time. If the number of concepts had gone up for each student that would imply that students were elaborating their mental models by thinking about more aspects of what it means to be a tutor. However, this is not what is happening. Indeed, for some students the number of concepts actually goes down. In contrast, for most students, density is higher at the end of the process (time 3) than at the beginning. This suggests that over time students are elaborating their mental models of what they want in a tutor by creating more intricate detailed models in which each of their ideas are better defined. Students are not adding requirements; rather, they are formulating their criteria and the inter-relationships among them more clearly.

For each student, there is a core set of concepts and statements that are used each time period (Table 7) Although students vary widely in the number of

ıt have happened", whereas
lid things backwards", "the
dirty".
ıll four students. In Fig. 6 a
:nts during the read-aloud
are shared by all students
:sent during the read-aloud
:ircled in Fig. 6. Note that
statements used by each
statements shared by all
: type of social knowledge
the nature of society. That
ıt "the boy has a mother"
: "boy has" and "boy got"
owledge about the physical
ons. Future work might see
; independent of the story

Table 7
Amount of information students hold constant over time

	Concepts	Statements	Density
Ian	17	43	2.53
Lowell	16	37	2.31
Jaques	22	52	2.36
Jubal	12	27	2.25
Hazel	24	62	2.58

Table 8
Amount of information shared by students at each time period

	Concepts	Statements	Density
Time 1	13	30	2.31
Time 2	17	43	2.53
Time 3	15	39	2.60
Across time	9	22	2.44

concepts and statements they use, they are remarkably similar in the density of these core maps. That is, if we consider just the concepts and statements that each student used in all three time periods, we see that for all students the density is between 2.25 and 2.58. This suggests that for each student there is a cognitive core description of what they want in a tutor that is consistent over time. Moreover, for all students, the level of complexity of this core is comparable. Whether what we are seeing is the result of some cognitive limitation or a coding artifact is a point for future research.

At any one time, across all students there exists a set of concepts and statements that all students share (Table 8). Neither the number of shared concepts nor statements increases consistently over time. However, the density does increase steadily. Across all three time periods there are a set of 9 concepts and 22 statements that all students share. These concepts and statements form the map shown in Fig. 7. Negative relations are indicated with dashed lines. We can think of this map as being the cultural core of the group. Here we see that the group of students all think of the tutor as someone who should be social. Being social means to fit in with the hall and to be friendly. Fitting in and being friendly means not insisting on quietness, or being provocative, being tolerant, reasonable, having the right personality, and interacting with other students. The fact that this core exists suggests that there is a minimal level of acceptability that is well agreed upon by the group. Outside of this level, however, is where differences in opinions begin.

5. Discussion

In this paper, both content analysis and map analysis has been used to explore various aspects of culture. Regardless of whether a content or a map analytic

Density
2.53
2.31
2.36
2.25
2.58

Density
2.31
2.53
2.60
2.44

similar in the density of
and statements that each
l students the density is
there is a cognitive core
over time. Moreover, for
rable. Whether what we
oding artifact is a point

of concepts and state-
r of shared concepts nor
e density does increase
of 9 concepts and 22
atements form the map
ied lines. We can think
/e see that the group of
cial. Being social means
:ing friendly means not
reasonable, having the
fact that this core exists
is well agreed upon by
nces in opinions begin.

s been used to explore
ent or a map analytic

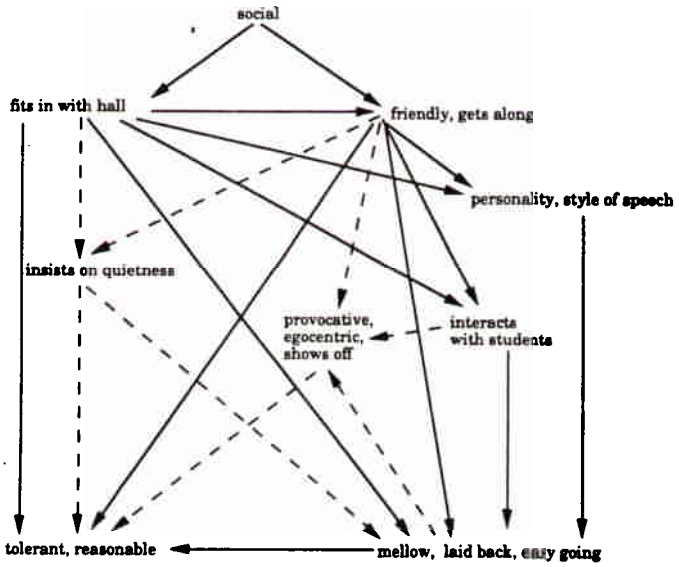


Fig. 7. Shared perception of tutor.

approach is taken doing textual analysis is non-trivial. That is, textual analysis requires a large amount of coding time. The content analysis techniques have an advantage, not discussed, and that is that they are more automated. This automation can substantially reduce the amount of coding time. The map analysis techniques are less automated. Indeed, future research should attend to developing more automated approaches to map analysis. In the meantime, the advantages of map analysis are that it has a theoretical basis and it allows the researcher to examine meaning.

Through analyzing texts, differences across cultures and changes within cultures can be located. Content analysis enables the researcher to locate the rhetoric of change and the extent to which different concepts are used. Map analysis takes the researcher a step further and enables the analysis of meaning. Four studies that illustrated the relative benefits of map analysis were described. These studies are extracted from larger more detailed studies. The reported findings should be viewed as illustrative and not conclusive. In all four cases, readers should turn to the original studies for additional details on the larger studies from which these illustrations are drawn. The studies reported in this paper are meant purely as illustrations of the kinds of cultural and cognitive analyses that are possible from both a content analytic and a map analytic perspective. In examining these studies the added value of map analysis for the cultural researcher can be seen.

In looking at the American perception of comedy and drama we saw that by using the thesaurus we could gain insight into general culture. In this analysis we saw that the added benefit of map analysis is that it made it possible to examine how different words, such as musical, could take on different meanings in different

contexts. Clearly this analysis did not give direct insight into human cognition. However, it would be possible to use these tools to examine how different individuals' mental models of comedy and drama differ from those 'classic' definitions in the thesaurus. Such a study could look at practitioners (actors, theater managers, etc.) and non-practitioners and provide insight into the extent to which this organizational field has an insider rhetoric.

In looking at the role of robots in science fiction we saw that by looking at the 'trend setters' we were able to observe cultural shifts over time in the way robots were thought of. In this analysis the added benefit of map analysis is that it made it possible for us to observe changes in the definition of robots. Map analysis also made it possible to see that while the definitions of robots became more diverse and more elaborate they did not continue to increase in density. In other words, the meanings of the words used to describe robots did not become more elaborate. The coded descriptions of robots are the authors' mental models. Although we did not examine this, it would have also been possible to look at the evolution of a single author's point of view. For example, we could contrast all the varying ways in which Asimov portrayed robots over time. Such an analysis would provide insight in to how an author's mental models change over time.

In looking at children's recall we saw/that there is a degradation in recall and that the recalled information is largely generic cultural information about people and their relationships. This suggests that story recall involves more than locating traditional story structures. It also involves reflecting on and restating cultural norms. In this analysis we saw that map analysis had the added benefit of making it possible to observe that the complexity of the recalled information is much simpler than that of the forgotten and generated information.

In looking at the students' maps of what they want in a tutor we saw that the students' perceptions shift noticeably over time. However, for each student individual (on average) and for the group as a whole the density of the maps increases over time. There is also a cultural core that is shared by all students across all time periods. This suggests that group decision making involves elaborating a common cultural core by explicating the meanings of the terms, and not simply adding more concepts.

These examples serve to illustrate the added benefits of map analysis. They demonstrate that by taking a more cognitive approach to coding texts we can gain additional information on culture. In general, it is often most informative to compare the results from a content analysis with those of a map analysis; e.g., to first examine concepts and then to move on to the examination of statements.

References

Berelson, B., 1952. Content analysis in communication research. New York: Hafner.
Blumer, H., 1969. Symbolic interactionism. Englewood Cliffs, NJ: Prentice Hall.
Carley, K., 1984. Consensus construction. Ph.D. Dissertation, Harvard University, Boston, MA.

Carley, K., 1986. An approach for relating social structure to cognitive structure. Journal of Mathematical Sociology 12(2), 137–189.
Carley, K., 1988. Formalizing the social expert's knowledge. Sociological Methods and Research 17, 165–232.
Carley, K., 1993. Coding choices for textual analysis: A comparison of content analysis and map analysis. In: P. Marsden (ed.), Sociological Methodology, Vol. 23. Oxford: Blackwell.
Carley, K. and D. Kaufer, 1993. Semantic connectivity: An approach for analyzing semantic networks. Communication Theory 3, 183–213.
Carley, K. and M. Palmquist, 1992. Extracting, representing and analyzing mental models. Social Forces 70, 601–636.
Cicourel, A. and K. Carley, 1990. The coder of narrative as expert: The ecological validity of coding practices. Presented at the American Anthropological Association Meetings.
Cicourel, A., K. Carley and M. Schratz, 1992. The interaction of cognition and cultural knowledge in coding children's recall of a story. Working Paper.
Danowski, J., 1980. Computer-mediated communication: A network-based content analysis using a CBBS conference. In: M. Burgoon (ed.), Communication Yearbook 6. Beverly Hills, CA: Sage.
Danowski, J., 1982. A network-based content analysis methodology for computer-mediated communication: An illustration with a computer bulletin board. In: R. Bostrom (ed.), Communication Yearbook. New Brunswick, NJ: Transaction Books.
Danowski, J., 1988. Organizational infographics and automated auditing: Using computers to unobtrusively gather and analyze communication. In: G. Goldhaber and G. Barnett (eds.), Handbook of Organizational Communication. Norwood, NJ: Ablex.
Ellis, F.S., 1968 [1892]. A lexical concordance to the poetical works of Percy Bysshe Shelley. New York: B. Franklin.
Ericsson, K.A. and H.A. Simon, 1984. Protocol analysis: Verbal reports as data. Cambridge, MA: MIT Press.
Fan, D., 1988. Predictions of public opinion from the mass media: Computer content analysis and mathematical modeling. New York: Greenwood Press.
Franzosi, R., 1990a. From words to numbers: A generalized and linguistic-based coding procedure for collecting textual data. In: C.C. Clogg (ed), Sociological methodology. San Francisco, CA: Jossey-Bass.
Franzosi, R., 1990b. Computer-assisted coding of textual data: An application to semantic grammars. Sociological Methods and Research 19, 225–257.
Garson, D., 1985. Content analysis machine-readable data file. V 1.0.
Gerbner, G. (ed)., 1969. The analysis of communication content; Developments in scientific theories and computer techniques. New York: Wiley.
Gottschalk, L.A. and G.C. Gleser, 1969. The measurement of psychological states through the content analysis of verbal behavior. Berkeley, CA: University of California Press.
Gottschalk, L.A. and C.N. Winget, 1969. Manual of instructions for using the Gottschalk-Gleser content analysis scales. Berkeley, CA: University of California Press.
Holsti, O.R., 1969. Content analysis for the social sciences and humanities. Reading, MA: Addison-Wesley.
Janowitz, M., 1969. Content analysis and the study of the 'symbolic environment'. In: A.A. Rogow (ed.), Politics, personality, and social science in the twentieth century. Chicago, IL: University of Chicago Press.
Kaufer, D. and K. Carley, 1993. Communication at a distance: The effect of print on socio-cultural organization and change. Hillsdale, NJ: Erlbaum.
Knorr-Cetina, K., 1981. The manufacture of knowledge. Oxford: Pergamon.
Latour, B. and S. Woolgar, 1979. Laboratory life. Beverly Hills, CA: Sage.
Mandler, J.M., 1984. Stories, scripts, and scenes. Aspects of schema theory. Hillsdale, NJ: Erlbaum.
Mead, G.H., 1962 [1934). Mind, self, and society. Chicago, IL: University of Chicago Press.
Mead, G.H., 1964. The genesis of the self and social control. In: A. Reck (ed.). New York: Bobbs-Merrill.
Namenwirth, J. and R. Weber, 1987. Dynamics of culture. Boston, MA: Allen and Unwin.

Neuman, W.R., 1989. Parallel content analysis: Old paradigms and new proposals. Public Communication and Behavior 2, 205–289.

North, R.C., 1963. Content analysis. A handbook with applications for the study of international crisis. Evanston, IL: Northwestern University Press.

Polanyi, M.P., 1958. Personel knowledge: Towards a post-critical philosophical. Chicago, IL: The University of Chicago Press.

Roberts, C.W., 1989. Other than counting words: A linguistic approach to content analysis. Social Forces 68, 147–177.

Stone, P.J., D.C. Dunphy, M.S. Smith and D.M. Ogilvie, 1968. The general inquirer: A computer approach to content analysis. Cambridge, MA: MIT Press.

Stryker, S., 1980. Symbolic interactionism. Menlo Park, CA: Benjamin Cummings.

Thomas, D., 1991. Changing cultural perception of robots as seen in science fiction. B.Sc. Thesis. Carnegie Mellon University.

VanLehn, K. and S. Garlick, 1987. Cirrus: An automated protocol analysis tool. In: P. Langley (ed), Proceedings of the Fourth International Workshop on Machine Learning. Los Altos, CA: Morgan-Kaufman.

VanMeter K.M. and L. Mounier, 1989. East meets West: Official biographies of members of the central committee of the communist party of the Soviet Union between 1981 and 1987, analyzed with Western social network analysis methods. Connections 1, 32–38.

Weber R., 1984. Computer-aided content analysis: A short primer. Qualitative Sociology 7, 126–147.

Young R., Stevenson W.B., Nicol T. and G.W. Gilmore, 1936. Analytical concordance to the Bible on an entirely new plan containing about 311,000 references, subdivided under the Hebrew and Greek originals, with the literal meaning and pronunciation of each designed for the simplest reader of the English Bible. New York: Funk and Wagnalls.