Using Representations to Assess Level of Membership in a Community of Practice

A Theory Paper Submitted to the CSCL '99 Workshop "Collaborative Use of Representations: Analyzing Learning Interactions"

CHRISTOPHER D. HUNDHAUSEN
Laboratory for Interactive Learning Technologies
Department of Information and Computer Sciences
University of Hawai`i at Manoa
Honolulu, HI 96822
hundhaus@hawaii.edu

Nearly every community of practice develops its own representations to support the practices of the community. For example, the community of electric circuit designers has developed circuit diagrams to represent electric circuits. Members of that community rely on such diagrams to discuss, communicate, and document what they do. Past research has established that members of a given community not only tend to interpret the community's representations in a similar way, but also tend to construct similar representations for a given problem or concept. I argue that the way in which an individual interprets and constructs the representations of a given community serves as a useful gauge of the individual's membership in the community. Drawing on *cultural consensus theory*, a framework developed within cognitive anthropology, I introduce an empirical method that uses the way in which individuals read and construct representations as a basis for substantiating the existence of a community, and as a basis for assessing an individual's level of membership in a community. To illustrate how the method can be applied, I present an example drawn from my own research into the use of graphical representations of algorithms ("algorithm visualizations") in computer science education. If one accepts the social constructivist premise that learning amounts to becoming a fuller member of a community of practice, then the method I introduce constitutes a richer and more valid means of measuring learning than traditional knowledge testing.

1 Introduction

EXTERNAL REPRESENTATIONS often play an instrumental role in the practices of a community. For example, electric circuit designers have devised specific ways of representing electric circuits, so that they can communicate their designs in a way that others in the community can understand. Air traffic controllers have become adept at reading radar plots of the local airspace, so that they can direct traffic appropriately. And in my own research, *algorithmaticians* (those who practice the design and analysis of computer algorithms) have developed standard ways of graphically representing an algorithm's procedural behavior, correctness, and efficiency, so that they can communicate about algorithms more effectively.

Recent empirical studies furnish evidence that the central members of a community read the representations of the community in a similar way. For example, Petre and Green (1993) describe a study that examined the graphical readership skills of expert and novice digital circuit designers. The participants were asked to answer a series of questions that required them to read and make inferences on textual and graphical representations of electronic circuits. A key result of the study was that the reading strategies of the experts "were more consistent as a group" (p. 63) than the reading strategies of the novices, which tended to be "chaotic" and "varied" (p. 63).

Likewise, recent empirical studies show that central members of a community tend to construct similar representations of a given problem or concept. For instance, Douglas, Hundhausen, and McKeown (1995; 1996) and Chaabouni (1996) present empirical studies of how humans construct visualizations of algorithms. In those studies, they asked computer science graduate students to construct visualizations of various sorting algorithms (bubble sort, insertion sort, heap sort, and quick sort) for the purpose of explaining the algorithms to a novice. An analysis of the underlying semantics of participants' visualizations revealed that, while they may differ significantly at a lexical level, their visualizations tended to portray an algorithm in terms of a similar set of semantic primitives.

Social constructivist learning theories such as the one proposed by Lave and Wenger (1991) view learning not in terms of knowledge acquisition, but rather in terms of increasingly central participation in the *practices of the community*. On this view, any observed differences in the ways in which various individuals perform community activities can be attributed to differences in the individual's *level of membership* in the community. A prime example of a significant community practice is the construction and meaningful interpretation of shared representations. Indeed, one would expect little variance in the ways in which the most central members of a given community of practice read and construct community representations. Conversely, one expect much divergence in the ways in which non-members of that community read and construct the community's representations.

The social constructivist position thus implies that an analysis of representations can serve as a useful basis for gauging community membership, and hence learning: In particular, an individual's level of membership in a given community can be gauged by that individual's *level of agreement* with other community members in the activities of *representation construction* and *representation interpretation*. Yet, how might we measure the level of agreement between two individuals with respect to how they interpret and construct community representations?

In this paper, I introduce *Cultural Consensus Theory*, a formalized, consensus-based model of community that has evolved out of research in cognitive anthropology. Using this theoretical framework as a guide, I develop an empirical method for measuring individuals' level of agreement on tasks that involve the interpretation and construction of community representations. I illustrate the application of the method by presenting an example from my own research into the use of graphical representations in computer science education. Many questions still remain with respect to the practicality and mechanics of the method—questions that only further empirical studies can answer. Nonetheless, I hope to make the case here that the method holds promise for social constructivist-minded educators looking for richer, more valid ways of measuring learning than traditional knowledge testing.

2 Cultural Consensus Theory

Over the past two decades, a line of research in cognitive anthropology has been concerned with formalizing a definition of culture based on *consensus*. The theory that has evolved out of that research—*Cultural Consensus Theory* (see, e.g., Boster, 1985; Romney, Weller, & Batchelder, 1986; Weller & Romney, 1988)—provides a suitable foundation both for determining whether a community of practice actually exists, and for assessing one's level of membership in such a community.

According to Consensus Theory, "each [community of practice] may be thought of as having an associated semantic domain that provides a way of classifying and talking about the elements in the culture pattern" (Romney, Weller, & Batchelder, 1986, p. 315). One can think of a semantic domain as an organized set of symbols (including notations, words, and graphical representations) for referring to a "common conceptual sphere" (Romney, Weller, & Batchelder, 1986, p. 315). For example, in the case of *algorithmaticians*, a common conceptual sphere might be a particular algorithm, and the organized set of symbols for referring to the algorithm might include a certain analysis involving Big-O notation; a certain correctness proof involving a loop invariant; a certain pseudocode description; a certain segment of source code in a programming language; and, most relevant to this research, certain graphical representations (algorithm visualizations). Moreover, Consensus Theory holds that the semantic domain associated with a given community of practice is a matter of *consensus*; the conceptual spheres characteristic of a community of practice, as well as the symbol systems appropriate for referring to them, are precisely those on which the members of the community agree.

Consensus Theory derives a formal statistical model for assessing the "cultural competence" of the individual members of a community of practice. The basis for such an assessment are the answers that a sample of community informants furnish to a set of questions. The questions must be carefully chosen so that they address a body of knowledge on which the community of practice is assumed to agree. However, unlike the statistical models traditionally applied to test-taking, Consensus Theory's statistical model does not assume an objective truth against which informants' answers are to be measured. Rather, the model uses informants' answers as a basis for constructing a *cultural truth*, according to which informants' cultural competence can then be assessed.[1]

Consensus Theory's statistical model constructs such a cultural truth based on patterns of agreement among informants. The assumption is that "the correspondence between the answers of any two informants is a function of the extent to which each is correlated with the [cultural] truth" (Romney, Weller, & Batchelder, 1986, p. 316). In other words, the most central members of the community, whose cultural knowledge is the most "complete," are highly likely to agree with each other by offering identical answers to the questions. In contrast, less central members of the community are less likely to agree both with each other, and with central members of the community.

A general strategy for performing a Consensus Analysis on a community of practice, then, involves three steps: first, developing a survey designed to elicit cultural knowledge that members of the community are assumed to share; second, administering that survey to a sample of people assumed to belong to the community; and third, analyzing informants' answers using Consensus Theory's statistical model.

3 Using Representation Interpretation and Construction as a Basis for Consensus Analysis

One of the pioneering applications of Consensus Theory aimed to account for observed differences in the way in which the South American Aguaruna Jivaro describe varieties of the *manioc* plant, their staple sustenance crop (Boster, 1985). Boster set up a garden containing 61 different varieties of the plant. One by one, he guided his 122 informants through this experimental garden, stopping at each plant to ask "What kind of manioc is this?" He recorded the words each informant used to describe the plants, and then performed a Consensus Analysis to determine and explore patterns of agreement among informants.

Obviously, the cultural activity around which Boster looked for consensus—*viz*., plant naming—is far simpler than the cultural activities of *representation interpretation* and *representation construction* that I am proposing to analyze using Consensus Analysis. Indeed, representation interpretation and representation construction involve far more than uttering words. It is not as though one can set up a "representation garden," guide informants through it, and ask them to name each representation. Rather, one must carefully design representation interpretation and representation construction tasks, so that informants' performances can ultimately be quantitatively evaluated for agreement.

In this section, I develop a method for using Consensus Analysis in order to calculate individuals' *cultural competence* with respect to the cultural activities of representation interpretation and representation construction. To illustrate the method, I present an example from my own research into the construction and interpretation of algorithm visualizations in computer science education. Although the method I present is geared specifically toward algorithm visualizations, I believe that it can serve as a foundation for carrying out a Consensus Analysis of the representation interpretation and construction activities of just about any community.

[1] Hence, the use of the term *informant* (as opposed to, say, *subject*) is deliberate; it underscores the fact that participants in a Consensus Study are *informing* the researcher of their culture, rather than the researcher *subjecting* them to a test.

3.1 Designing the Consensus Survey

The consensus "survey" that I propose engages informants in the two activities that are of interest here:

(1) *representation interpretation*—meaningfully interpreting a community representation; and

(2) *representation construction*—constructing a representation that other community members will understand.

In my own research, I have designed tasks involving the interpretation and construction of *algorithm visualizations*. My interest in these tasks was spurred by ethnographic studies I carried out in an actual undergraduate algorithms course in which students constructed their own visualizations (Hundhausen, 1999, ch. 4). In those studies, I observed that the ability to interpret and to construct graphical representations of algorithms was shared by *schooled algorithmaticians*—those who participate in the practice of designing and analyzing computer algorithms. Through constructing and presenting their own visualizations in course assignments, and through viewing visualizations constructed by others, students gradually developed an ability to interpret and to construct graphical representations of algorithms—an ability that, by the end of the course, more closely resembled the interpretation and construction skills of their instructor. In other words, by the end of the course, upon watching a particular visualization, students increasingly agreed on how it mapped back to the algorithm it represented. Similarly, by the end of the course, students began to graphically represent the same algorithm in terms of a similar set of semantic elements. In short, one measure of students' learning in the course was their ability to interpret and construct the representations of the community that is in the process of reproducing itself through an undergraduate algorithms course: the *Community of Schooled Algorithmaticans.*

3.1.1 The Interpretation Task

In the interpretation task I propose, informants are presented with a well-established community representation. They are then asked to map graphical elements, attributes, and transformations of that representation to their underlying semantics. For example, members of the Community of Schooled Algorithmaticians would be asked to view an identical, well-established visualization of a well-known algorithm. Figure 1 presents a well-established "sticks" visualization of *bubblesort*, a well-known sorting algorithm whose pseudocode is presented in Figure 2. In this visualization, the sticks represent the elements to be sorted; stick height represents element magnitude. As the visualization progresses, adjacent stick elements are flashed, and smoothly switch places if they are out of order. By the end of the visualization, stick elements form an ascending or descending ramp, indicating that they are in order.

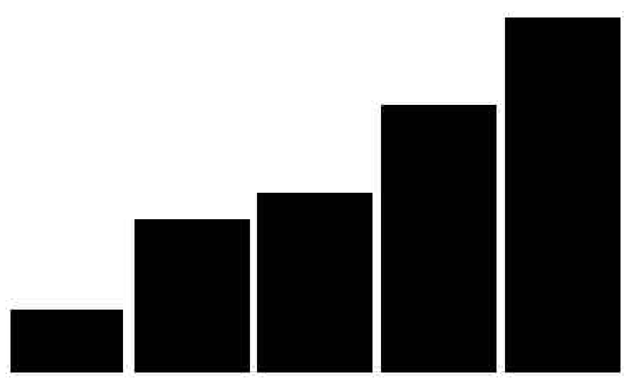

Figure 1. Classic "Sticks" Visualization of the Bubblesort Algorithm

```
BUBBLE-SORT(A)
for j ← n - 1 to 1 do
  for i ← 1 to j do
    if (A[i] > A[i+1] then
      exchange A[i] ↔ A[i+1]
    end if;
  end for;
end for;
end BUBBLE-SORT;
```

Figure 2. Pseudocode Description of the Bubblesort Algorithm

To elicit informants' interpretation of this visualization, we would ask informants to build a table mapping significant elements, attributes, and transformations in the visualization to their underlying meanings. So that participants have a common basis for their mappings, we would ask them to anchor the meanings in a common pseudocode description of the algorithm, such as the one presented in Figure 2. Table 1 presents an example of what a table constructed by an informant might look like. Stick elements are mapped to entries in the pseudocode array "a." Stick height (an attribute of stick element) is mapped to values of the pseudocode array "a." The visualization transformation in which two elements flash, as well as the visualization transformation in which two elements switch places, is mapped to a specific block of pseudocode.

SEMANTICS	OBJECT/ATTRIBUTE/TRANSFORMATION
Element of array A	Stick
Value of element of array A	Height of stick
array "a"	Row of sticks
if (A[i] > A[i+1] **then**	Adjacent sticks flash
exchange A[i] ↔ A[i+1]	Adjacent sticks change places

Table 1. Mapping the Bubblesort Visualization Shown in Figure 1 to its Semantics

As part of my dissertation research, I have conducted pilot studies that included a task like this one. I observed that informants who view an identical visualization can and will construct different sets of mappings. This observation underscores an important point implied by Consensus Theory: namely, that community representations *do not* derive their meaning not from absolute rules of human perception or graphic presentation. Rather, the meaning of community representations, just like the meaning of natural languages, is a matter of agreement among members of the community.

3.1.2 The Construction Task

I propose a construction task in which informants are first presented with a problem or concept that is of interest to the community. They are then asked to construct their own representations of that problem or concept. For example, members of the Community of Schooled Algorithmaticians would be given a pseudocode description of a well-known algorithm, for which they would be asked to construct a visualization. In prior empirical studies (Douglas, Hundhausen, & McKeown, 1995; Douglas, Hundhausen, & McKeown, 1996), we asked computer science graduate students to construct visualizations of the bubblesort algorithm (described in the previous section) for the purposes of explaining the algorithm to an introductory-level computer science student who is unfamiliar with the algorithm. Participants used simple art supplies (colored pens, colored construction paper, scissors) to construct their visualizations. They "executed" their homemade visualizations by doing such things as (a) moving construction paper "cutouts" across the table, (b) gesturing, (c) marking up their visualizations dynamically with a pen, and (d) providing a play-by-play narrative of the action as it unfolded. Figure 3 presents snapshots of the three visualizations that we observed in the study.

On the surface, the visualizations presented in Figure 3 appear to vary widely. However, by conducting the same kind of semantic-level analysis that we would have informants perform in the interpretation task described above, we can map these visualizations to their underlying semantics, as shown in Tables 2 and 3. Table 2 maps the objects of the visualizations, along with the objects' attributes, to their underlying semantics. Table 3 maps significant visualization transformations to lines of pseudocode. As these tables indicate, even though the three visualizations differ widely at a lexical level, they all portray the bubblesort algorithm in terms of a similar semantics. In the next section, I illustrate that it is possible to use this kind of semantic level analysis as input to a Consensus Analysis.

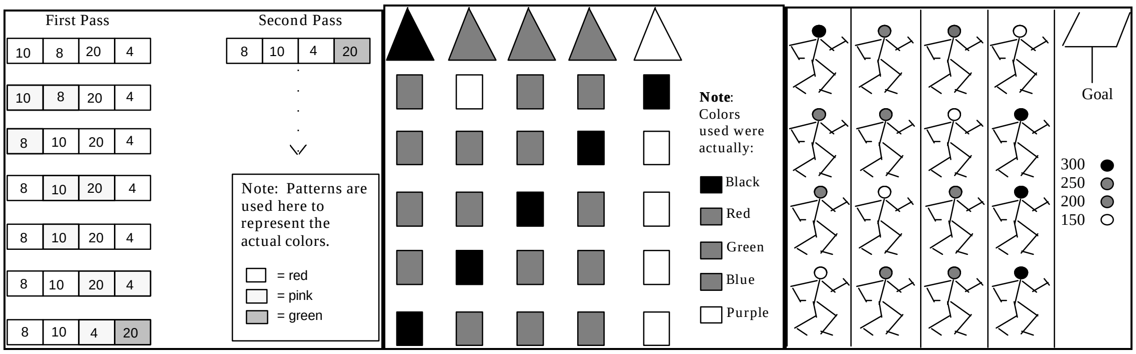

Figure 3. Snapshots of the Number, Color and Football Visualizations of the Bubblesort Algorithm Observed in the Studies of Douglas, Hundhausen, & McKeown (1995, 1996)

SEMANTICS	NUMBER AV LEXICON	COLOR AV LEXICON	FOOTBALL AV LEXICON
entry in array "a"	Square	Square	Stick Figure
value of entry in array "a"	Number symbol	Color	Color (as weight)
array "a"	Contiguous row of squares	Non-contiguous row of squares	Contiguous row of figures
Inner loop pass history	Rows of sorting elements	—	—
Outer loop pass history	Columns of rows	Rows of sorting elements	Rows of sorting elements
Legend explicating ordering on sort elements	—	Triangles with color spectrum	Column of color/player weight pairs

Table 2. Mapping the Lexical Objects and Attributes of the Bubblesort Visualizations Observed by Douglas, Hundhausen, & McKeown (1995, 1996) to Their Semantics

SEMANTICS	NUMBER AV LEXICON	COLOR AV LEXICON	FOOTBALL AV LEXICON
DO outer loop	Start new column of rows	Create new row of squares	Create new row of football players
DO inner loop	Create new row of squares	—	—
a) Reference elements to be compared	Color elements pink	—	Location of football
b) Compare elements (same, <, >)	—	—	Intuitions about how player size relates to running, tackling, and fumbling
c) Exchange elements	Exchange numbers	Exchange colors	Ball carrier advances by tackling next football player in line (thereby exchanging positions with that player)
d) Don't exchange elements	—	—	Fumble football to next player in line
Terminate outer loop	Color square in correct order green	—	—
Terminate Sorting	Ordering of natural numbers, all squares green	Color squares match legend	Players ordered by weight

Table 3. Mapping the Lexical Transformations of the Bubblesort Visualizations Observed by Douglas, Hundhausen, & McKeown (1995, 1996) to Their Semantics

3.2 Quantifying Agreement

In order to use Consensus Theory's statistical model to assess an informant's "cultural competence," we first need a means of quantifying the extent go which two individuals agree. Otherwise, we have no way of determining the consensus of a community around the activities of representation interpretation and construction.

To determine the level of agreement between two individuals with respect to their performance in the representation interpretation and construction tasks just outlined, we can make use of the semantic-level analysis technique just described. With respect to representation interpretation, two informants are said to agree to the extent that they can view a representation and perform a similar semantic-level analysis on the representation. More formally, let s_1 be the set of semantic primitives (i.e., lexical-to-semantic mappings) that informant i_1 gleans from viewing a representation, and let s_2 be the set of semantic primitives that informant i_2 gleans from viewing the same representation. Then the *proportion of agreement* between informants i_1 and i_2 can be defined as the proportion of semantic primitives in s_1 and s_2 that are identical:

$$\textit{Proportion of agreement}(i_1,i_2) = \frac{|s_1 \cap s_2|}{|s_1 \cup s_2|} \qquad \text{(Eq. 1)}$$

In the case of representation construction, the researcher, not the informant, is performing the semantic level analysis. Other than this difference, the agreement between two informants can be calculated in an analogous fashion: Two informants are said to agree to the extent that the representations that they construct have semantic primitives in common. More formally, let s_1 be the set of semantic primitives determined to exist in the representation of informant i_1 for problem or concept p, and let s_s be the set of semantic primitives determined to exist in the representation constructed by informant i_2 for the same algorithm or concept p. Then Eq. 1 expresses the proportion of agreement between the representations of i_1 and i_2.

To illustrate how one might apply this definition of agreement, Table 4 computes the pair-wise levels of agreement among the three visualizations presented in Figure 3, based on the semantic-level analysis presented in Tables 2 and 3.

PAIR	PROPORTION OF AGREEMENT (OBJECTS AND ATTRIBUTES)	PROPORTION OF AGREEMENT (TRANSFORMATIONS)	PROPORTION OF AGREEMENT (TOTAL)
Number-Color	4/6 (67%)	3/6 (50%)	7/12 (58%)
Number-Football	4/6 (67%)	4/8 (50%)	8/14 (57%)
Color-Football	5/6 (83%)	3/6 (50%)	8/12 (67%)

Table 4. Computing the Pair-Wise Proportion of Agreement among the Visualizations Presented in Figure 3

This definition of informant agreement raises three general issues that require further discussion. First, it is unclear whether analysts will use similar vocabulary to describe mappings. If they do not, it may be difficult to interpret whether their analyses actually concur. Based on pilot studies I have conducted to explore these issues, I have concluded that analysts choices on both sides of the mapping must be sufficiently *constrained*, so that agreement between them is not subject to interpretation, but rather can be determined unambiguously. Constraining analysts' choices entails requiring them to identify unambiguously each visualization element they wish to map (perhaps by having them use software to point-and-click at each element), and requiring them to express semantics in terms of a common pseudocode description.

A second issue raised by the above definition of agreement is that, in the case of the representation construction task, the sets s_1 and s_1 are subject to the *interpretation* of an analyst trained in the semantic-level analysis technique. To minimize the bias of a single analyst, we would ideally have a set of trained analysts perform independent semantic-level analyses on informants' representations;

the *intraclass reliability coefficient* (see, e.g., Shrout & Fleiss, 1979) of the analysts could then be incorporated into the Consensus Analysis.

The third issue raised by the above definition of agreement is that it may oversimplify the notion of semantic agreement. Indeed, Tversky (1977) has developed a formal model for assessing the agreement between two sets of semantic features that is far more sophisticated than Eq. 1 (set intersection divided by set union). In particular, given a domain of feature (i.e. semantic primitive) sets $\{s_1, s_2, \ldots, s_n\}$, Tversky defines a function $\mathbf{s}(a,b)$ to be the relative similarity between feature sets a and b. The function $\mathbf{s}(a,b)$ assigns a value to each pair of feature sets, such that $\mathbf{s}(a,b) > \mathbf{s}(c,d)$ means that a is more similar to b than c is to d. The function $\mathbf{s}(a,b)$ is defined as a *linear contrast* between the semantic primitives in set a and b:

$$\mathbf{s}(\mathrm{a,b}) = \theta\mathbf{f}(a \cap b) - \alpha\mathbf{f}(a - b) - \beta\mathbf{f}(b - a) \qquad \text{(Eq. 2)}$$

where $\mathbf{f}(a \cap b)$ counts the number of primitives that a and b have in common; $\mathbf{f}(a - b)$ counts the number of primitives belonging to a but not to b; $\mathbf{f}(b - a)$ counts the number of primitives belonging to b but not to a; and θ, α, and β are weightings indicating the relative importance of these three entities. Notice that, in contrast to Eq. 1, Eq. 2 provides a framework for weighting the relative importance of various features of a semantic domain. In performing a Consensus Analysis, one would do well at least to experiment with the Tverskian model, in addition to using the intuitive formula for proportion of agreement presented in Eq. 1.

3.3 A Sample Consensus Study

Based on the ideas just discussed, I conclude this section by sketching out an example of out how one might design and analyze a Consensus Study of representations. As discussed above, my own research is interested in the practices of the community in the process of reproduction through an undergraduate algorithms course—what I call the “Community of Schooled Algorithmaticians.” This community shares a sense of not only what is important and interesting about algorithms, but also of how to communicate about algorithms. Such communication almost invariably involves the use of visual representations of algorithms—“algorithm visualizations.” Thus, for social constructivist-minded educators, students’ convergence on expert-like performance in visualization interpretation and construction constitutes an important gauge of learning within an undergraduate algorithms course.

A Consensus Study centered around these activities could be used to measure the effectiveness of alternative pedagogical treatments in facilitating students’ convergence on expert competence in an undergraduate algorithms course. In order to establish the “cultural truth” in these tasks, we must first have a group of expert algorithmaticians (i.e., instructors) complete visualization interpretation and construction tasks (see Section 3.1) for two different algorithms. The results of our semantic level analysis of their tasks would be fed to Consensus Theory’s statistical model, in order to determine the “cultural truth” for these tasks—that is, the "culturally correct" set of semantic primitives that experts glean from (*interpretation* tasks) the visualizations that they view, and include in (*construction* tasks) the visualizations that they construct in the tasks.[2]

Once we run the experts through our study, we are ready to recruit our student informants. The idea is to employ a between-subjects design in which we divide students into two groups, each of which completes an alternative set of learning exercises. Before completing these learning exercises

[2]According to Consensus Theory’s statistical model, the higher the average cultural competence among informants, the fewer informants are required to determine the “cultural truth” with a high level of confidence. For example, if the average level of cultural competence among participants is 0.8 (i.e., informants agree with each other 80% of the time), then we need only 6 participants to determine the “cultural truth” at the .95 confidence level (Romney, Weller, & Batchelder, 1986). Thus, the number of expert informants we would actually need to attain an acceptably high level of confidence (0.95) cannot be known precisely in advance. Rather, it is determined by the results of intermediate Consensus Analyses. A preliminary estimate is that six to ten experts would be needed.

(which may span just a few hours, but more likely will span an entire academic semester), both groups complete interpretation and construction tasks for the first of the two algorithms. The results of a semantic-level analysis of these tasks are fed into a Consensus Analysis, so that we can determine students' baseline cultural competence with respect to the expert informants in the study. Then, upon completing their respective pedagogical treatments, both groups of student informants complete interpretation and construction tasks for the second of the two algorithms. Once again, the results of a semantic-level analysis of these tasks are fed into a Consensus Analysis, in order to determine students' "final" cultural competence with respect to the expert informants in the study.

At this point, we are in a position to test for significant differences between the two groups. If the average change in one group's level of cultural competence (that is, the difference between its final and baseline cultural competence) is significantly higher than the average change in the other group's level of cultural competence, then that would be grounds for concluding that one pedagogical treatment is more effective than the other.

4 Conclusion

This paper has introduced a novel method for measuring an individual's competence within the important community activities of constructing and meaningfully interpreting external representations. The premise, rooted in sociocultural constructivist learning theory (Lave & Wenger, 1991), is that learning amounts to increasingly central participation in a community of practice—that is, an ability to perform activities that are important to the community in a way that increasingly resembles a community expert's performance. The theoretical foundation of the method is *Consensus Theory*, a formalized, consensus-based model of community that has evolved out of research in cognitive anthropology. According to Cultural Consensus Theory, there is no such thing as absolute truth. Rather, Consensus Theory's formal model makes it possible to derive a "cultural truth" based on patterns of agreement among community members.

The main obstacle to using Consensus Theory as a means of measuring individual competence with respect to representation construction and interpretation is the difficulty in quantifying the extent to which two individuals agree in these activities. To overcome this obstacle, I have introduced a technique called *semantic-level analysis*, in which the lexical features of a representation are mapped to their underlying semantics. The number of semantic elements that representations have in common can then be used as a basis for quantifying the extent to which they are in agreement. Such a quantitative measure of pair-wise agreement is all that Consensus Theory's model requires in order to derive the *cultural truth* around representation construction and interpretation, and in order to determine each individual's level of cultural competence.

Through an example from my own research into the use of algorithm visualization in computer science education, I have demonstrated how one might use the method introduced here as a novel means of measuring learning—one that incorporates learners' performance in constructing and interpreting the representations of a community of practice. I hope that the example I have presented has convinced you of the potential of this method. At the same time, I want to underscore that what I have presented here is just a beginning. Consensus Theory originally evolved out of an interest in studying differences in the names people have for things such as plants. Shifting from the relatively simple domain of names to the potentially complex domain of graphical representations is quite a leap. Clearly, much empirical research remains to be done in order to refine the method, and in order to demonstrate its feasibility and utility in practice.

References

Boster, J. S. (1985). "Requiem for the omniscient informant": There's live in the old girl yet. In J. W. D. Dougherty (Ed.), *Directions in Cognitive Anthropology* (pp. 177-197). Urbana: University of Illinois Press.

Chaabouni, Z. D. (1996). A user-centered design of a visualization language for sorting algorithms. Master's Thesis, Department of Computer and Information Science, University of Oregon, Eugene, OR.

Douglas, S. A., Hundhausen, C. D., & McKeown, D. (1995). Toward empirically-based software visualization languages. In *Proceedings of the 11th IEEE Symposium on Visual Languages* (pp. 342-349). Los Alamitos, CA: IEEE Computer Society Press.

Douglas, S. A., Hundhausen, C. D., & McKeown, D. (1996). Exploring human visualization of computer algorithms. In *Proceedings 1996 Graphics Interface Conference* (pp. 9-16). Toronto, CA: Canadian Graphics Society.

Hundhausen, C. D. (1999). Toward Effective Algorithm Visualization Artifacts: Designing for Participation and Communication in an Undergraduate Algorithms Course. Ph.D. Dissertation, Department of Computer and Information Science, University of Oregon, Eugene, OR.

Lave, J., & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. New York: Cambridge University Press.

Petre, M., & Green, T. R. G. (1993). Learning to read graphics: Some evidence that 'seeing' an information display is an acquired skill. *Journal of Visual Languages and Computing 4*, 55-70.

Romney, A. K., Weller, S. C., & Batchelder, W. H. (1986). Culture as consensus: A theory of culture and informant accuracy. *American Anthropologist 88*(2), 313-338.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin 36*(2), 420-428.

Tversky, A. (1977). Features of similarity. *Psychological Review 84*, 327-352.

Weller, S. C., & Romney, A. K. (1988). *Systematic Data Collection*. Newbury Park, CA: Sage.

Brief Author Bio

Christopher Hundhausen has actively pursued research into the use of algorithm visualization technology in computer science education ever since he was a college freshman at Lawrence University, from which he graduated with a B.A. in math/computer science in 1991. While at Lawrence, he assisted Thomas Naps in the development of an algorithm visualization laboratory for four summers under an NSF grant. After beginning graduate work at the University of Oregon, Chris became increasingly interested in the human side of algorithm visualization technology, as he began asking questions like "Why are we building this technology, if we have no evidence that it actually assists in learning?" Under the guidance of his mentor, Sarah Douglas, he pursued coursework and independent study projects that reflected this emerging interest in human-computer interaction, and graduated with an M.S. in computer science in 1993.

In 1993, Chris was awarded a Fulbright grant to pursue his algorithm visualization research at the University of Karlsruhe in Germany, under the mentorship of Walter Tichy. While in Germany, Chris led a research group that carried out empirical research into how humans specify algorithm visualizations, and he taught a graduate seminar on user-centered software development.

In the fall of 1994, Chris entered the Ph.D. program at University of Oregon, where he began his current research into the use of external representations in general, and algorithm visualizations in particular, within communities of practice. Chris's dissertation research, supervised by Sarah Douglas, blended techniques from a variety of disciplines —anthropology, psychology, education, and computer science—in an attempt to build an empirical case for the pedagogical use of algorithm visualization technology. He graduated from the University of Oregon with a Ph.D. in June of 1999. Visit his dissertation web site (http://lilt.ics.hawaii.edu/~hundhaus/dis) for further information.

After getting married to Laura Girardeau in July of 1999, he moved to Hawaii, where he presently works (when he isn't at the beach or in the rainforest) as a postdoctoral associate for Daniel Suthers in the Laboratory for Interactive Learning Technologies at the University of Hawai`i at Manoa.