Cultural Consensus Analysis as a Tool for Clinic Improvements

C. Scott Smith, MD, Magdalena Morris, RN, MS, CS, William Hill, PhD, Chris Francovich, EdD, Juliet McMullin, PhD, Leo Chavez, PhD, Caroline Rhoads, MD

Some problems in clinic function recur because of unexpected value differences between patients, faculty, and residents. Cultural consensus analysis (CCA) is a method used by anthropologists to identify groups with shared values. After conducting an ethnographic study and using focus groups, we developed and validated a CCA tool for use in clinics. Using this instrument, we identified distinct groups with 6 important value differences between those groups. An analysis of these value differences suggested specific and pragmatic interventions to improve clinic functioning. The instrument has also performed well in preliminary tests at another clinic.

KEY WORDS: **patient-centered care; organizational cultures; education; medical.**
J GEN INTERN MED 2004;19:514–518.

Clinics may face recurring problems that arise when participants have differing values. Physicians value evidence and logic, while patients focus on understanding and choice.[1] Patients want to be heard,[1] while residents want to efficiently move to the next patient.[2] Faculty and resident values differ for "medical facts" and euthanasia.[3,4] It is not known to what extent these value differences lead to recurring problems. Evaluating these differences is difficult, and the inability to do so may hinder attempts to improve processes.

Cultural consensus analysis (CCA) comes from the field of anthropology.[5] It is theoretically based on Kroeber's discussion of "systemic culture patterns" of language and meaning.[6] CCA assumes that cultural knowledge is shared and systematically distributed.[5,6] Group affiliation may be inferred by similarity of response to a set of meaningful statements. Important differences in cultural knowledge and values between groups may be elucidated by differing responses to the same set of statements. This method has been used to better understand cancer screening[7,8] and to more accurately target screening interventions.[9] We were interested in the possibility of using CCA to identify "operational subcultures" or groups defined by their role in clinic (e.g., patient, faculty, resident) that may have conflicting values. These data could illuminate opportunities for improvement in clinic operations. No one has used CCA in this fashion.

Received from the Northwest Regional Faculty Development Center (CSS, MM, WH, CF), VA Medical Center, Boise, Idaho; Departments of Medicine (CSS, CR), Medical Education and Biomedical Informatics (CSS), University of Washington, Seattle, Wash; Department of Nursing (MM), Boise State University, Boise, Idaho; and Departments of Medicine (JM) and Anthropology (LC), University of California, Irvine, Calif.

Presented in part at the Fifth Congress of the International Society for Cultural Research and Activity Theory (ISCRAT), June 18–22, 2002, Amsterdam, The Netherlands, and the 26th Annual Session of the Society of General Internal Medicine (SGIM), May 1–3, 2003, Vancouver, Canada.

Address correspondence and requests for reprints to Dr. Smith: Medicine Service (111), 500 W. Fort Street, Boise, ID 83702 (e-mail:scott.smith2@med.va.gov).

METHODS

This study was part of a larger ethnographic study of resident learning at the VA Medical Center in Boise, Idaho. The methods have been discussed in more detail elsewhere.[10] The University of Washington human subjects division and the institution's research and development committee approved this study. Development of the CCA was carried out in 5 steps, as follows.

Ethnography

Two trained observers (MM and another research nurse) performed over 130 hours of observation and interviews of patients and staff in clinic and created a detailed description of the clinic. Two separate analysts (CSS and CF) used a mix of free coding (themes that emerged from the data) and template coding (an outline of likely-to-be-important categories) to recursively categorize these observations. Discrepancies were resolved by discussion and revision of the category definition if necessary. A stable set of categories was identified, and they were combined into plausible conceptual models for each group based on agreement with the observational data. The credibility of these conceptual models was then confirmed by checking them with each group. Tensions between groups were predicted from these models, and these were used to create focus group questions.

Focus Groups

Trained facilitators held 2 focus groups each for patients, faculty, and residents (6 in all). These were audiotaped and transcribed. The trigger questions were:

- In general, what do you want from a clinic visit?
- In what ways do you feel that patients and doctors want the same or different things from a clinic visit?
- How do you feel about the resident teaching program at Boise?
- How does changing your doctor affect your care (of the patient)?
- How does the computer affect the clinic visit?
- In general, how much time should a clinic visit take?

Two anthropologists (JM and LC) who were blinded to the ethnography and conceptual models extracted the data.

In order for CCA to have the power to differentiate between groups and yet not be overwhelming, 10 to 20 different statements are ideal. Our consultants selected the 16 most salient responses, based on frequency of occurrence, for use in our CCA. These were transcribed, one to a card, onto 16 laminated 3×5 inch cards.

Statement Revision

The Flesh-Kincaid reading level of the original instructions and cards was 10th grade, analyzed using Microsoft Word. We were able to reduce this to just above 5th grade without affecting meaning. We asked several subjects to rank order the cards and then explain their choices. Most statements showed a clear and consistent interpretation among and between groups. A few showed bimodal response patterns. These were explained by differences in interpretation. For instance, many patients interpreted "Having enough staff around to help…" to mean "staff helping doctors." However, other patients interpreted this as "staff helping patients." The wording of these ambiguous statements was clarified.

Performance of the CCA and Validation

Convenience samples of all subjects were approached during prespecified time periods for verbal consent. CCA statements (Table 1) on cards were presented to 20 patients, 10 faculty (67% of total at site), and 10 residents (77% of total at site). The 16 cards were spread out randomly in a circle on a flat surface. Subjects were asked to pick up the cards in order of importance from most to least important. Sorting order was recorded on a standardized form.

In the analysis, we assumed 1) there is a set of cultural values that each group shares, 2) each subject is more or less familiar with this set (cultural knowledge; C), and 3) the more matches that two subjects share, the higher their C values (in a precise mathematical relationship).[5] Briefly, a factor analysis is performed much like psychometricians would in test construction, only applying it to informants rather than items see appendix for more detail.[11] This provides an estimate of C for each subject. The "culturally correct" answers, which would align that subject with his cultural partners, are then calculated *a posteriori* using Bayes's theorem.[5] The standard for assuming a shared cultural model is a ratio of ≥3:1 between the eigenvalues of the first and second factor.[5] Analysis was performed in Anthropac (Anthropac 4.0, Analytic Technologies, Columbia, SC).

With this method, the different rank orders of the statements by patients, faculty, and residents were compared. The 6 largest differences between groups were hypothesized to cause significant recurring problems in clinic. This was confirmed by comparison with observational, interview, and focus group data.

Generalizability

Harborview Medical Center, a county hospital located 600 miles away, was chosen as a second site. The CCA cards were again presented to 20 patients, 10 faculty, and 10 residents using convenience sampling. No subjects overlapped between the two sites. The order of statements was recorded and analyzed as described above. A correlation matrix between patients, faculty, and residents from both sites was created.

Table 1. Statements Used for Cultural Consensus Analysis, Average Group Ranking of the Statement, and Maximal Difference Between Group Rankings

Statement	Patients' Ranking	Faculty Ranking	Residents' Ranking	Maximal Difference
Have the same doctor for more than 1 year	1	11	12	11
Use a computer to check the patient record	14	10	6	8
See the patient within 15 minutes of the appointment time	7	15	10	8
Talk to the patient about healthy lifestyle changes (such as exercise, stop smoking, limit alcohol)	12	5	5	7
Senior doctor reviews student doctor's work	9	13	16	7
Doctor gets a reminder to talk about healthy habits and testing for silent diseases	15	9	13	6
Doctor asks what is changing in the patient's life (such as a move or major family changes)	11	6	9	5
Have senior doctors around to answer questions for student doctors	10	12	15	5
Doctor and patient agree on goals	6	3	3	3
Stay on time to see as many patients as possible	13	14	11	3
Let the patient know about the lab results	5	7	7	2
Dictate the clinic note	16	16	14	2
Take the time to find the cause of the pain or sickness the patient is feeling	2	1	1	1
Talk to patients until they understand what the doctor is doing	3	2	2	1
Get quick treatment for the pain or sickness the patient is feeling	4	4	4	0
Have enough people around to help the doctor with telephone calls, blood work, and shots	8	8	8	0

RESULTS

Ethnography

Each group had a different conceptual metaphor for the clinic visit. Credibility of these models was high when checked with members. They saw themselves in their own model, and confirmed the other group's behaviors accordingly. For patients, the model was "a series of locks" (check in, wait, vital signs, wait…). For residents, it was a "docket" model (who controls what is on the agenda today). For faculty, it was a "balance" model (ideal vs realistic).

Focus Groups

The tones of focus groups for patients, faculty, and residents were different. Patient groups mentioned (often very emotionally) the difficulty in getting what they needed and how dehumanizing and inefficient the clinic could be. They were careful not to criticize specific elements that they were dependent upon (e.g., their own doctor).

Resident groups focused on efficiency. The doctor-patient relationship was described abstractly. Resident groups exhibited normative process, tending toward agreement and wanting to appear "role appropriate."

Faculty groups were rebellious and challenging, and focused on specific doctor-patient relationships. The original 16 statements for CCA cards were extracted from transcripts of these focus groups.

Performance of the CCA and Validation

The aggregate of patients, faculty, and residents did not share a cultural model by standard assumptions (eigenvalue ratio must be ≥3; for the aggregate it was 1.43).[5] Patients demonstrated near cultural consensus (ratio, 2.56), faculty showed strong cultural consensus (ratio, 3.84), and residents showed very strong cultural consensus (ratio, 6.07). This suggests that values about clinic are strongly tied to role for residents, patients are more multidimensional, and faculty fall in between.

When the average rank order of the statements by these 3 groups was compared, there were 6 high-difference statements (Table 1) that represented 68% of the variance. These differences correlated with problems identified by observational, interview, and focus group data. For instance, the difference in CCA rankings for the "computer" statement did seem to reflect important value differences between groups. It agreed well with focus group data.

> "How does the computer affect the clinic visit?"
> Patient focus group: "I don't like it. Some of the things they say, I'm constantly filling in the blanks."
> Faculty focus group: "Getting data out is a value. Putting data in is a pain in the neck."
> Resident focus group: "I can always get whatever it is that I want."

It also agreed with direct observations.

> Resident: (typing and clicking on the computer)
> Patient: (begins to say something and then stops midword; notices the resident is dealing with the computer) "Excuse me."
> Resident: (doesn't notice the interruption)
> Patient: (flushes and looks irritated)

Finally, it agreed with interview data.

> Patient: (to researcher) "Now that [doctors] use those computers, they ask me how I'm doing; I tell them and they say, 'It doesn't say that in here!' They just don't listen to me!"

Generalizability

Results of this CCA at Harborview were similar to Boise. Patients, faculty, and residents did not share a cultural model. Their strength of agreement was in the same order, with very nearly the same eigenvalue ratios. Five of the six statements with the largest between-group differences were the same. The sixth, about computers, was explained by very different computer support at the two sites. Finally, as seen in Table 2, correlations between the same groups at different hospitals (e.g., Boise faculty-Harborview faculty) were higher than correlations between different groups at the same hospital (e.g., Boise faculty-Boise residents).

DISCUSSION

To our knowledge, this is the first study to use CCA for detecting important value differences in groups defined by their role in clinic. Our CCA instrument was able to differentiate between patients, faculty, and residents and to

Table 2. Correlation Matrix for the CCA Rankings of Patients, Residents, and Faculty from the Boise VAMC and Harborview Medical Center (HMC)

	Pt-Boise	Res-Boise	Fac-Boise	Pt-HMC	Res-HMC	Fac-HMC
Pt-Boise	1.00					
Res-Boise	0.04	1.00				
Fac-Boise	0.18	0.70	1.00			
Pt-HMC	0.87	0.12	0.19	1.00		
Res-HMC	0.06	0.84	0.55	0.58	1.00	
Fac-HMC	0.19	0.43	0.82	0.66	0.76	1.00

identify large value differences between these groups that correspond to problems documented by observation and interview. Initial tests at one outside institution look promising for generalizability.

Weaknesses of this study include the fact that it was conducted primarily at a single institution. Validation was retrospective. Nonetheless, the data provided new insights into clinic problems and has catalyzed quality interventions, as discussed below.

Cultural consensus analysis used in this way may be a promising improvement tool. Relatively small sample sizes are sufficient when there is modest average sharing of a single cultural model.[5] CCA has credibility with constituents that other data may not: it is based in a ground-up approach supported by ethnography; the conceptual models and responses "ring true" to experience; it is based on explicit mathematical criteria; and comparing the rankings between groups often provides new insights. As used in this study, however, CCA is resource intensive, which may limit its wide applicability as a practical quality improvement tool.

It is not known whether a CCA instrument created in one clinic can be generalized to others. This should be further studied. CCA assumes each respondent is independent, so they must be interviewed individually.[5] It also assumes all statements are drawn from one "domain" or area of shared culture. A CCA created in one clinic would only be expected to provide meaningful results in similar clinics.

As an example of the utility of CCA data, the leadership in our training program reviewed these results. They were surprised and disturbed by the large value discrepancy with regard to "same doctor for more than one year." When several residents and faculty were questioned directly, they believed that continuity was important but they could not see a way to avoid turnover in a training program or to increase faculty continuity without undermining the resident's role. The magnitude of the discrepancy inspired a brainstorm for solutions including:

- Dedicating a "no-patients" clinic at the beginning and end of each resident's tenure so a faculty supervisor can "sign in" and "sign out" the clinic panel (like on-service, off-service notes).
- New patients enter the system through a resident panel and stay a maximum of 3 years, then bump up to a faculty panel.

These questions revealed tacit values (productivity, politics of access) that were limiting our thinking. Both of these improvements are now being implemented.

The relationships between other large intergroup differences in shared cultural values (CCA differences) and major recurring problems in a given clinic are not well understood. For instance, in our data, a large CCA difference in the value of computers was dependent on the extent of computerization (Boise vs Harborview). This may require special training in doctor-patient-computer skills at sites with computerized medical records. These relationships should be studied further at multiple sites.

This material is based upon work supported by the Office of Research and Development, Health Services R&D Service, Department of Veterans Affairs (grant PCC 98-010).

REFERENCES

1. Laine C, Davidoff F, Lewis CE, et al. Important elements of outpatient care: a comparison of patients' and physicians opinions. Ann Intern Med. 1996;125:640–5.
2. Mizrahi T. Getting Rid of Patients: Contradictions in the Socialization of Physicians. New Brunswick, NJ: Rutgers University Press; 1986.
3. White JA, Anderson P. Learning by internal medicine residents: differences and similarities of perceptions by residents and faculty. J Gen Intern Med. 1995;10:126–32.
4. Bushwick B, Emrhein D, Peters K. A comparison of resident and faculty attitudes toward physician-assisted suicide and active voluntary euthanasia. Fam Med. 2000;32:261–6.
5. Romney AK, Weller SC, Batchelder WH. Cultural consensus: a theory of culture and informant accuracy. Am Anthropol. 1988; 88:313–37.
6. Kroeber AL. Anthropology. New York: Harcourt, Brace; 1948.
7. Chavez LR, Hubbell FA, McMullin JM, Martinez RG, Mishra SI. Understanding knowledge and attitudes about breast cancer. A cultural analysis. Arch Fam Med 1995;4:145–52.
8. McMullin JM, Chavez LR, Hubbell FA. Knowledge, power and experience: variation in physicians' perceptions of breast cancer risk factors. Med Anthropol 1996;16:295–317.
9. Hubbell FA, Chavez LR, Mishra SI, Magana JR, Burciaga Valdez R. From ethnography to intervention: developing a breast cancer control program for Latinas. J Natl Cancer Inst Monogr. 1995; 18:109–15.
10. Smith CS, Morris M, Francovich C, et al. A qualitative study of resident learning in ambulatory clinic. The importance of exposure to "breakdown" in settings that support effective response. Adv Health Sci Educ. 2004; in press.
11. Comfey AL. The minimum residual method of factor analysis. Psychol Rep. 1962;11:15–8.

APPENDIX A

Cultural Consensus Analysis Calculation

The mathematical analysis assumes that participants share a cultural model, causing their rankings to be similar. Let C_1 equal participant 1's cultural knowledge—the probability of producing the "culturally correct" (and unknown to us) answer. Let N equal the number of statements being sorted. Then the probability (P_1) that participant 1 ranks the group's most important response first is:

1. $$P_1 = C_1 + (1 - C_1)/N$$
(knows) (guesses)

Rearranging this equation, the first participant's competence is:

2. $$C_1 = (N\,P_1 - 1)/(N - 1)$$

The chance that participant 1 and participant 2 both rank a statement first is the sum of them both knowing, C_1C_2,

plus one knowing and two guessing, $C_1(1 - C_2/N)$, plus one guessing and two knowing, $C_2(1 - C_1/N)$, plus both guessing $(1 - C_1)(1 - C_2)/N$, equals:

3. $$P_{1,2} = C_1C_2 + C_1(1 - C_2/N) + C_2(1 - C_1/N) + (1 - C_1)(1 - C_2)/N \quad \text{or}$$

4. $$P_{1,2} = C_1C_2 + (1 - C_1C_2)/N$$

Thus, the agreement between any two participants is a function of the product of their cultural knowledge. On average, the more they agree, the greater their cultural knowledge. Rearranging the above equation:

5. $$C_1C_2 = (N\,P_{1,2} - 1)/(N - 1)$$

This key relationship, that the product of any two participants' cultural knowledge is a function of their similarities of response corrected for guessing, is then utilized.

Equation 5 is one equation with two unknowns (C_1 and C_2), and cannot be solved. However, an $S \times S$ matrix (S, number of participants) of matches between participants contains $S \cdot (S - 1)/2$ point estimates with S unknowns, and can be analyzed by "best fit." A least squares factor analysis of this matrix, using the minimum residual method,[11] estimates the cultural knowledge (C) for each participant. If the assumption of shared cultural knowledge is true, then it should explain the greatest amount of variance in the data. The standard assumption is an eigenvalue ratio of ≥3:1 between the first (cultural knowledge) and second (any other) factor. This acts as a check on the assumptions of the model.

If shared culture is established by this criterion, the *a priori* estimates of C, and the actual rankings by each individual, are then used to determine *a posteriori* the "culturally correct" answers for each cultural group.[5] These rankings are then compared between groups, looking for important discrepancies.