USING PSYCHOLOGICAL SCALING TECHNIQUES TO ASSESS CLINICAL EXPERTISE IN ANESTHESIOLOGY

Olena Connor
Department of Psychology
New Mexico State University
Las Cruces, New Mexico

Nancy J. Cooke
Applied Psychology
Arizona State University East
Mesa, Arizona

Matthew B. Weinger
Department of Anesthesiology
University of California San Diego and the
Anesthesia Ergonomics Research
Laboratory
VA San Diego Healthcare System
San Diego, California

Jason Slagle
Department of Anesthesiology,
University of California San Diego and the
Anesthesia Ergonomics Research Laboratory
San Diego Healthcare System
San Diego, California

This study used Pathfinder, a psychological scaling technique, to assess underlying cognitive structure associated with mastery of relevant knowledge necessary for anesthesiology decision-making. Our study revealed this approach to be a valid method to assess the tacit knowledge that underlies clinical expertise. A set of concepts associated with a decision-to-extubate scenario was derived from expert interviews. Participants included nine attending anesthesiologists, seven first-year anesthesiology residents, and eight second-year anesthesiology residents. Pathfinder was applied to participants' pairwise relatedness judgments of the clinical concepts in the context of the scenario. Experts' data were aggregated to form an expert referent structure. Student anesthesiologists were assessed based on comparison of their structures to this referent. These comparisons yielded a knowledge score that was highly correlated with residents' exam grades. This finding supports our position that Pathfinder is a valid knowledge assessment method and, as a complement to current exams, can be applied to assess a student's deep understanding of anesthesiology concepts.

INTRODUCTION

Studies on clinical expertise, employing behavioral task analysis and workload assessment during actual patient care, have demonstrated that more experienced anesthesia providers have different clinical task distributions, perform clinical tasks more efficiently, are more vigilant, have lower workload, and are under less stress than novice residents (Weinger et al, 1994, Weinger & Slagle, 2001). While such approaches document *what* clinicians actually do during patient care, they do not elicit information about *why* they do it (i.e., the underlying knowledge structures involved in clinical decision-making). This information will guide not only clinician training and evaluation, but the design and evaluation of medical devices and care processes.

This study used Pathfinder, a conceptually-oriented psychological scaling technique for knowledge elicitation, to assess the knowledge structure underlying anesthesiology expertise (Cooke, 1994). Pathfinder generates networks from conceptual proximity judgments, where concepts are represented as nodes and their interrelations as links between nodes (Schvaneveldt, 1990). These judgments are assumed to reflect the judge's underlying semantic organization that is associated with nonmonotonic mastery of the relevant medical knowledge (Patel, Arocha, & Kaufman, 1991). In comparison to traditional written or oral examinations, which require direct recall of information, psychological scaling techniques are applied to relatively indirect conceptual proximity judgments. These judgments are indirect in that they only require an overall assessment of relatedness with no explanation for, or description of, the nature of that relation. Thus, knowledge that is relatively implicit can be conveyed through these judgments in the same way that partial or implicit knowledge may be better reflected in a multiple choice test than an essay test. However, in comparison to multiple-choice tests that only provide an overall assessment of amount of information acquired, psychological scaling can provide this assessment, plus a way to infer a rich conceptual structure from the set of judgments (Goldsmith, Johnson, & Acton, 1991).

Pathfinder has been successfully used to assess expertise in a variety of domains such as air combat flight maneuvers (Schvaneveldt, Durso, Goldsmith, et al., 1985) and computer programming (Cooke & Schvaneveldt, 1988). However, this approach has been only sporadically applied to medicine (McGaghie, Boerger, McCrimmon, & Ravitch, 1994). Even so, Pathfinder has not been applied to high-tempo procedurally oriented acute care specialties like anesthesiology where expertise may not be as readily discerned with other assessment techniques (such as interviews, observations, etc.). Moreover, many previous medical applications have lacked the context specificity that is critical for unambiguous participant judgments and thus, successful assessment (Chase & Simon, 1973; McGaghie, McCrimmon, Mitchell, Thompson, & Ravitch, 2000).

We hypothesize that Pathfinder networks, based on judgments of the anesthesiology concepts grounded in context of a scenario, will reveal greater differences in cognitive structure between groups of participants who differ in level of experience than within groups of individuals of similar level of experience. Moreover, we hypothesize that a comparison between novice and experts knowledge will correspond to exam performance. This approach is proposed to be a valid method to assess the tacit knowledge that that underlies clinical expertise, and may prove to be complementary to other evaluation techniques.

METHOD

Design

In this study level of experience in anesthesiology (experts who were attending anesthesiologists or novices who were first- and second-year resident anesthesiologists) was the predictor variable with pairwise conceptual relatedness ratings and their Pathfinder representation as the dependent measure.

Participants

Twenty-four participants from University of California-San Diego varied in their level of clinical anesthesiology expertise and included: seven first-year anesthesiology residents, eight second-year anesthesiology residents, and nine attending anesthesiologists. All participants were contacted via e-mail or telephone. Experts and residents, who were available and who agreed to participate in this study (that is 24 out of those 27 who were asked to participate), were paid \$10-\$25 per hour depending on position.

Materials

Based on structured interviews with expert anesthesiologists, a realistic scenario that involved a difficult extubation decision and 16 clinical concepts relevant to this case were created. Concepts were chosen for which pairwise relatedness ratings were expected to discriminate experts from novices. These concepts can be seen in Figure 1 and the scenario is presented in the Box 1 below.

Box 1. Clinical Scenario that Provided Context for Relatedness Ratings.

Clinical Scenario:

You are called urgently to OR 4 to take over an almost completed case so that the anesthesia provider can immediately go to the emergency room to assess a critically ill patient requiring truly emergent surgery. The emergency surgery is scheduled to immediately follow the case you are taking over because this is the only available OR. The departing anesthesiologist rapidly provides the following information about the on-going anesthetic:

<u>Preoperative Assessment:</u>

Mr. Jones is a 42 year-old male who weighed 66 kg and was 5'6" tall. He presented with right sciatica and lower leg weakness and was scheduled for an L_4-L_6 with instrumentation and bone graft. His past medical history was notable only for a one-pack per day smoking history. He had not had any prior surgical procedure and was not allergic to any medications. Mr. Jones took 1-2 Vicodin every 6 hours for chronic back pain. His preoperative blood pressure was 135/65 torr, heart rate 76 beats/min, respiratory rate 16 breaths/min and temperature 36.8°C. His room air oxygen saturation was 98%. The physical exam was unremarkable – he had a normal airway exam. Mr. Jones had been NPO since midnight and the evaluating anesthesiologist preoperatively assigned him an ASA 2 physical status.

<u>Intraoperative Course:</u>

After the institution of routine monitors and preoxygenation, Mr. Jones was induced with a typical dose of propofol. Intubation was easily accomplished after a single dose of succinylcholine. Arterial and foley catheters were inserted prior to placement in the prone position. Anesthesia was maintained with isoflurane (1-1.5%) and nitrous oxide 60% in oxygen as well as intermittent doses of fentanyl. Throughout the case, he had received a total of 20 µg/kg of fentanyl with the last dose given 30 min prior to end of surgery. Muscle relaxation had been maintained throughout the case with pancuronium, a total of 12 mg, with the last dose administered 30 min prior to end of surgery.

During the 5.5 hour procedure, the patient lost 1800 ml of blood and excreted 450 ml of urine. Fluid replacement included 3 units autologous blood and 4500 ml crystalloid. The final hematocrit was 30. There was one twitch noted of a train of four with the nerve stimulator on the ulnar nerve and this was reversed with 4 mg of neostigmine and 0.8 mg of glycopyrolate. The patient was turned to the supine position without incident.

You assume care of the patient 15 minutes after the end of surgery and 10 minutes after termination of anesthesia. The current situation is …

Mr. Jones is opening his eyes spontaneously and beginning to move his arms toward his face. The nerve stimulator revealed a full train of four and a sustained tetanus. He is breathing at a rate of 23 breaths/minute with a tidal volume of 300 ml. The end-tidal CO_2 is 52 torr and the end-tidal isoflurane concentration was 0.25%. The other vital signs are blood pressure 167/87 torr, heart rate 98 beats/min, temperature 35.8°C, and the oxygen saturation on 100% inspired oxygen is 97%. The next patient is on her way up from the ER and the nurses are already trying to quickly turn over the room.You must decide whether to extubate the patient <u>now</u> or to take him to the PACU first and extubate him at a later time.

Procedure

Each participant was presented with the scenario and complete list of concepts in order to give the rater the idea of the scope of the to-be-rated concepts. The concept pairs were displayed on a Macintosh laptop one-at-a-time (i.e., *total dose of narcotic administered – surgical duration*). Order of the items in the pairs was counterbalanced across participants and pairs were displayed in a random order (without replacement). Participants were asked to judge the relatedness of each concept pair in the context of the scenario on a 5-point scale (5 = slightly related, 1 = highly related) with a discrete sixth point for unrelated concept pairs.

Data Analysis

The rating procedure resulted in a 16 x 16 symmetrical matrix of proximities (values of 1-6) for each participant. Matrices were submitted to Pathfinder network scaling, a data reduction routine that eliminates direct links if a shorter or equivalent indirect path exists. Pathfinder analysis was performed using the KNOT software tool (Schvaneveldt, 1990) with default parameter settings. The output from this method is a network structure with concepts represented as nodes and strong relations as links between nodes. In addition, aggregate networks were also generated for each group of participants by including those links that were present in a majority of the group member networks. Links in these aggregate networks are weighted by the number or proportion of group members whose networks included that link. Networks of two or more individuals or groups differ in terms of the presence or absence of links. The KNOT tool also generates a quantitative value (C-value) that reflects the similarity between two networks based on the proportion of shared links.

RESULTS

Group Cohesiveness

Intragroup agreement on pairwise relatedness ratings of domain concepts is typically greater for experts than novices (Cooke & Schvaneveldt, 1988). Within and between groups' pairwise correlations of relatedness ratings revealed less intragroup similarity among experts than was expected (See Table 1). Variance within the expert group was greater than anticipated possibly due to sizeable variance in post-residency experience (3-28 years).

Table 1. Average Within and Between Groups Pairwise Correlations of Relatedness Ratings.

	Expert (9 experts)	2nd Year (n=8)	1st Year (n=7)
Expert	.50**	.48**	.43**
2nd Year	--	.44**	.41**
1st Year	--	--	.40**

Note: **Correlation is significant at the 0.01 level (2 tailed)

A follow-up cluster analysis of experts' relatedness ratings indicated two outliers whose ratings were the most dissimilar from the other experts. These individuals considerably were also more similar to each other in regard to their experience and their demographics. Those two attending anesthesiologists were excluded from the expert group. New analyses yielded higher intra-group similarity for the seven experts (r = .56) compared to the original correlation of .50. The difference between the original expert correlation of .50 and the new expert correlation of .56 is statistically significant (t = 2.00, p= .03). Pairwise correlations between this reduced set of experts versus first-year residents and second-year residents resulted in average correlations of .47 and .49 respectively.

The aggregate network based on links in the majority of the seven attending anesthesiologists' networks was used as an expert standard or "referent" against which each student was compared. C-values between each resident's network and the "referent" showed the degree of similarity in knowledge organization between experts and each novice. For example, the network of one of the second year residents was the most similar to the expert referent (C-value is .43), whereas a first-year resident's network was the least similar to this "referent" (C-value is .16). C-values between the referent and first-year residents (mean = .30) were slightly lower than C-values between the referent and second year residents (mean = .32). However, a t-test indicated that this difference was not statistically significant, t = 2.18, p = .72.

The comparison between the "referent" and each resident's network constitutes a knowledge score. As a test of the validity of our Pathfinder-based approach, this knowledge score was further correlated with oral exam grades for each of the second year residents. The oral exam focused primarily on clinical decision making (in the context of case-based scenario), but also covered basic anesthesiology knowledge. Two independent examiners alternated in questioning each resident for approximately 35 minutes and rated their answers. A thirds observer, who did not communicate with residents, also rated the answers. The residents also did rating self-assessment before getting a feedback from the faculty. The correlation between the results from this clinical exam and the Pathfinder derived knowledge score was marginally significant, r =.70, p =.06. This finding supports the validity of the Pathfinder approach as an indicator of residents' knowledge.

Qualitative Comparisons

To explore qualitative differences between the three clinician groups, an aggregate network was generated for each group (the two outlying experts were excluded from the expert network). Figure 1 displays a comparison between the expert and first-year aggregate networks.

As was predicted by the three of the four experts who participated in pre-data collection interviews, the "*Oxygen Saturation*" concept did not differentiate experts and novices. The majority of experts, as well as the majority of novices, did not see a critical connection between the *Oxygen Saturation* concept and other concepts with regard to decision making in the context of the presented clinical scenario.

Further, it was predicted that the importance of the concept "*Residual Anesthetic Effect*" would be underestimated by novices, whereas the experts would see it as quite important. Indeed, the experts viewed this concept as highly central, connecting it to five other concepts in the consensus network. Yet, the novices did not appear to appreciably underestimate this concept, connecting it to four other concepts. On the other hand, we can readily see

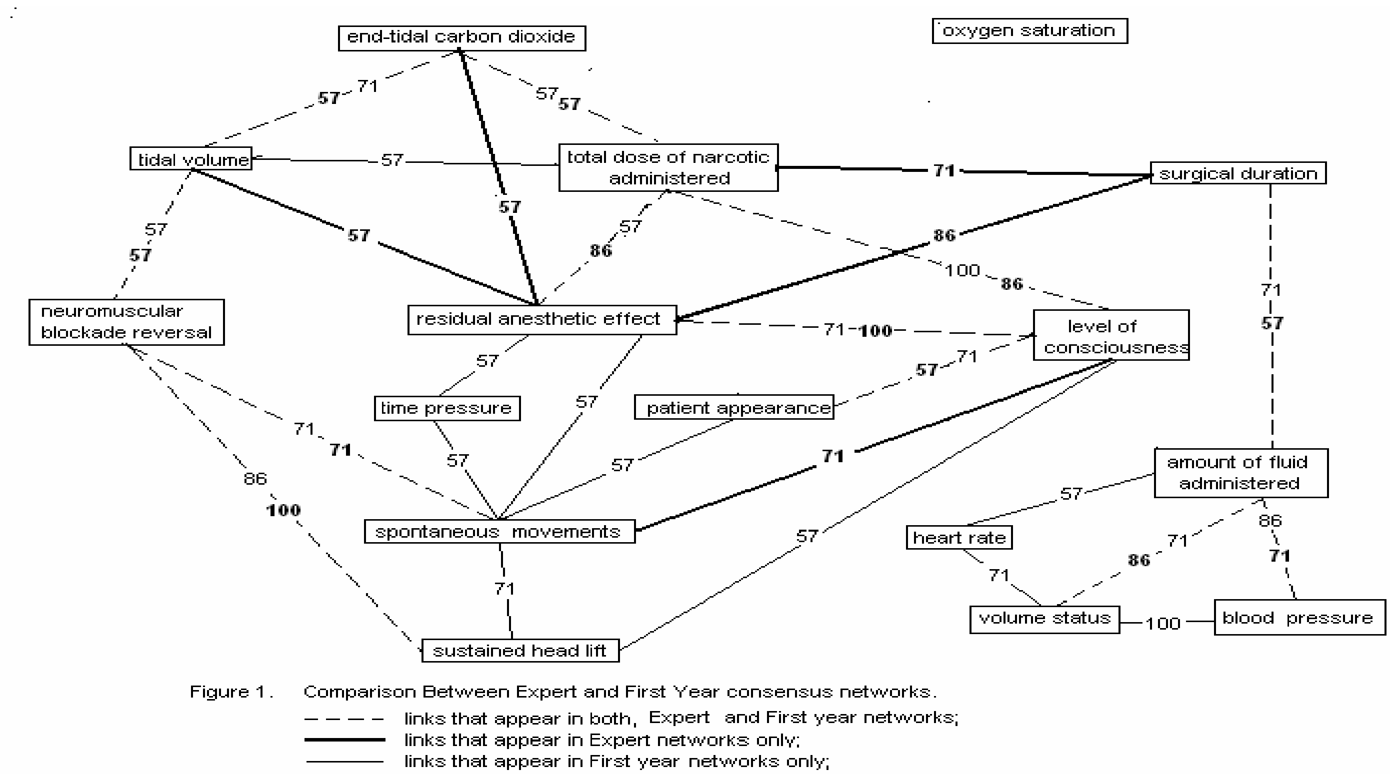

Figure 1. Expert versus first-year aggregate networks. The values shown on the links indicate the percentage of individual expert networks (presented in bold font) and individual first-year networks (presented in normal font) that contained that link.

at least one specific expert-novice difference; 86% of experts connected *Residual Anesthetic effect* to "*Surgical Duration*", whereas novices did not see this connection whereas experts saw a direct connection to only two other concepts. Again, it is interesting to note the specific differences between the two groups. Seventy-one percent of experts connected *Spontaneous Movements* to "*Level of Consciousness,*" whereas, novices did not see this as a strong relationship.

Table 2 summarizes some of the critical concept pairs (represented by shadowed cells) that may differentiate experts from second-year residents and first-year residents.

Table 2. Critical concept pairs that differentiate experts from novices. The percentage of the seven experts with that link is indicated in bold font. Second-year resident percentages are in italic font and first-year percentages are in normal font.

CONCEPT PAIRS	Tidal Volume	Surgical Duration	End-tidal Carbon Dioxide	Level of Conscious-ness	Volume Status	Time Pressure	Sustained Head Lift	Heart Rate
Residual Anesthetic Effect	**57**, *75*, 0	**86**, *63*, 0	**57**, *75*, 0	**100**, *88*, 71	**0**, *0*, 0	**0**, *63*, 57	**0**, *63*, 0	**0**, *50*, 0
Total Dose of Narcotic Administered	**0**, *0*, 71	**71**, *50*, 0	**86**, *75*, 57	**86**, *50*, 100	**0**, *0*, 0	**0**, *0*, 0	**0**, *0*, 0	**0**, *0*, 0
Amount of Fluid Administered	**0**, *0*, 0	**57**, *0*, 71	**0**, *0*, 0	**0**, *0*, 0	**0**, *86*, 71	**0**, *50*, 0	**0**, *0*, 0	**0**, *57*, 0
Spontaneous Movements	**0**, *0*, 0	**0**, *0*, 0	**0**, *0*, 0	**71**, *50*, 0	**0**, *0*, 0	**0**, *0*, 57	**0**, *88*, 71	**0**, *0*, 0

DISCUSSION

In this study we investigated a method for obtaining a conceptual representation of concepts relevant to decision making by anesthesiologists. Most researchers agree that the development of expertise is characterized not so much by further knowledge acquisition, but by reorganizing the knowledge in a way so that it is more readily available (Chi, Feltovich, & Glaser, 1981). Written or oral examinations are limited in their ability to assess usable clinical knowledge, focusing on amount of information acquired as opposed of the practical organization of that information. Moreover, direct classroom tests do not adequately tap tacit knowledge associated with level of expertise. Further, since knowledge organization influences decision making performance (Gaba, 1992), it is critical to assess knowledge structure for different levels of expertise. Revealing students' knowledge organization may help instructors understand students' misconceptions or faulty mental models depending on how dissimilar a student's knowledge network is from the expert "referent". The innovation of this study lies in its attempt to assess level of anesthesiology expertise indirectly by applying psychological scaling techniques to clinical views of the strength of relatedness of concept pairs in the context of a specific scenario.

Obtained results revealed different cognitive structures depending on the level of the clinical expertise. It was predicted, the differences in cognitive structure between groups of participants who differed in level of experience would be greater than within groups of individuals with similar experience. The quantitative within and between group comparisons did not show any significant difference. However the qualitative comparisons of the network structures highlighted some noticeably differences.

Our second hypothesis addressing the validity of the psychological scaling method was supported by high correlations with oral exam scores. Thus, this method could can be very useful in medical education in addition to oral exams, experts' evaluations and self-evaluations, because the latter methods, especially oral exams which are greatly influenced by communication skills, are more subjective in their ability to assess deep, tacit knowledge.

Moreover, applying this method in conjunction with other knowledge assessment methods will allow for richer and deeper understanding of how clinical material is conceptually organized by anesthesiology students and targeting of potential misconceptions. Also, the indirect knowledge assessment methods may be applied as alternative educational exams a few times per semester, as this would be cheaper and less time consuming relative to oral exams. Not only will this methods aid in assessing ongoing knowledge acquisition with the final goal of student's network approaching the expert network referent, but it will also help instructors to identify weak areas in the student's knowledge structure.

REFERENCES

Chi M, Feltovich PJ, Glaser R. Categorization and representation of physics problems by experts and novices. *Cognitive Science.* 1981;5:121-152.

Chase, W.G., & Simon, H.A. (1973). Perception in chess. *Cognitive Psychology*, *4*, 55-81.

Cooke, N. J. (1994). Varieties of knowledge elicitation techniques. *International Journal of Man-Machine Studies*, *41*, 801-849.

Cooke, N. J., & Schvaneveldt, R. W. (1988). Effects of computer programming experience on network representations of abstract programming concepts. *International Journal of Man-Machine Studies*, *29*, 407-427.

Gaba, D. M. (1992). Improving anesthesiologists' performance by simulating reality. *Anesthesiology, 76,*491-494.

Goldsmith, T. E., Johnson, P. J., & Acton, W. H. (1991). Assessing structural knowledge. *Journal of Educational Psychology*, *83*, 88-96.

McGaghie, W. C., Boerger, R. L., McCrimmon, D. R., & Ravitch, M. M. (1994). Agreement among medical experts about the structure of concepts in pulmonary physiology. *Academic Medicine*, *69*, S78-80.

McGaghie, W. C., McCrimmon, D. R., Mitchell, G., Thompson, J. A., & Ravitch M. M. (2000). Quantitative concepts mapping in pulmonary physiology: Comparison of student and faculty knowledge structures. *Advances in Physiology Education*, *23*, 72-81.

Patel, V. L., Arocha, J. F., & Kaufman, D. R. (1991). Expertise and tacit knowledge in medicine. *Tacit Knowledge in Professional Practice: Practitioner Perspective* (pp. 75-99). Mahwah, NJ: Lawrence Erlbaum Associates.

Schvaneveldt, R. W. (1990). *Pathfinder Associative Networks: Studies in Knowledge Organization*. Norwood, NJ: Ablex.

Schvaneveldt, R. W., Durso, F. T., Goldsmith, T. E., Breen, T. J., & Cooke, N. J. (1985). Measuring the structure of expertise. *International Journal of Man-Machine Studies*, *23*, 699-728.

Weinger, M. B., Herndon, O. W., Paulus, M. P, Gaba, D., Zornow, M. H., & Dallen, L. D. (1994). Objective task analysis and workload assessment of anesthesia providers. *Anesthesiology* 80: 77-92.

Weinger, M. B. & Slagle, J. M. (2001). Task and workload analysis of the clinical performance of anesthesia residents with different levels of experience (abstract). *Anesthesiology*. *95*, A1145.