Prototypes and semantic qualifiers: from past to present

Georges Bordage

Clinical reasoning has been the focus of a rich body of scholarship in medical education, for which the work of Arthur Elstein and his colleagues, on the Medical Inquiry Project at Michigan State University, has served as an early platform. A major thread in this research on clinical reasoning concerns our investigations of prototypes and semantic networks, conducted over the last several years. This essay reviews the development and main results of our work and formulates 3 main lessons gleaned from it, related to the dynamic interaction among theories, methods and practice, the importance of theory testing and theory building, and the versatile nature of knowledge organisation in memory.

According to structural semantics, knowledge is given meaning through networks of relationships represented by dichotomous abstract qualifiers

One of the main conclusions of the work by Elstein *et al.* was already becoming apparent when I first met Arthur Elstein in 1975. It was later reported in their book, *Medical Problem Solving: an Analysis of Clinical Reasoning*: 'The difference between experts and weaker problem solvers are more to be found in the repertory of their experiences, organised in long-term memory, than in differences in the planning and problem-solving heuristics employed.'[1] A couple of years later, this conclusion led me to ask how medical knowledge is organised in memory and to test an accepted theory of knowledge organisation at the time, prototype theory (led by the work of Rosch and Mervis[2–4]), as it might apply to medical knowledge. According to prototype theory, medical categories would be organised in memory around representative exemplars, the prototypes that serve as anchors for other members of the category. For example, gastrointestinal disorders would be organised around duodenal ulcers, gastritis and Crohn's disease. Other more peripheral members of the category might include cirrhosis, malabsorption, Meckel's diverticulitis, ileus and colon cancer. I conducted 4 studies to gather converging evidence for the presence of prototypes in the memory of medical students and experienced family doctors. The 4 studies consisted of:

1 a free recall task of disorders belonging to 14 categories of medical knowledge (i.e. organ systems [e.g. respiratory, endocrine, musculoskeletal], pathophysiological mechanisms [e.g. inflammation] and major patient complaints [e.g. abdominal pain]);

2 a judgement task measuring the degree of representativeness of the category members from study 1;

3 a free recall task of the attributes of the same disorders, and

4 a measure of response time for prototypical (central) versus peripheral members as measured in study 2.

The results from these 4 studies provided converging evidence for the presence of prototypical disorders in the memory of medical students and experienced doctors.[5] The disorders recalled first were judged to be more representative than those recalled later, had more attributes in common (family resemblance), and were recalled faster and more accurately than the peripheral members.

Successful diagnosticians use semantic qualifiers more frequently and in more diversified sets in their discourses than diagnosticians who are less successful

Subsequently, 2 education corollaries from prototype theory were tested using observational studies, namely, that prototype formation was related to medical school courses that contained fewer disorders ($r = -0.58$) and more intermediate level disorders ($r = +0.73$; e.g. angina pectoris compared with coronary insufficiency, a superordinate category, or Prinzmetal angina, a subordinate category).[6] This reflected a strategy policy of 'less is more' in teaching and learning prototypes, where the initial presentation was limited to 4 or 5 main causes of shoulder pain, for example, rather than all 32 causes in the textbook. Prototype formation was also fostered by presenting intermediate level disorders to students. Osler had understood and applied these

Department of Medical Education, College of Medicine, University of Illinois at Chicago, Chicago, Illinois, USA

Correspondence: Dr Georges Bordage, Department of Medical Education (M/C 591), College of Medicine, University of Illinois at Chicago, 808 South Wood, Chicago, Illinois 60612-7309, USA. Tel: 00 1 312 996 7349; Fax: 00 1 312 413 2048; E-mail: bordage@uic.edu

doi: 10.1111/j.1365-2923.2007.02919.x

principles a century before by dwelling on 2 disorders, pneumonia and typhoid, with the clerks working on his ward because '…if thoroughly understood by the students, [pneumonia and typhoid] give them a satisfactory foundation on which to build their later experience' (in Cushing[7]). He also encouraged his peers to teach less, for 'the student tries to learn too much, and we the teachers try to teach him too much – neither, perhaps, with great success'.[8]

Theory, practice and research methods constantly interact to provide a dynamic interplay that helps move theory and practice forward

By the mid-1980s, the questions I was asking had evolved and now focused on the nature of the relationships that link pieces of knowledge in memory. Prototype theory did not provide an adequate conceptual framework with which to address this new question, but structural semantics provided a promising theoretical basis. According to structural semantics, knowledge is given meaning through networks of relationships represented by dichotomous abstract qualifiers (axes), such as disorders associated with acute versus chronic onsets or local versus systemic manifestations. For example, 4 sets of semantic qualifiers (axes) go a long way towards providing a mental scaffolding for knowledge about low back pain, that is, acute–chronic, immediate–delayed, above–below the knee and local–systemic. Thus, when an otherwise healthy man complains of a lower back pain that began the previous day when he lifted a heavy object and that ran down his right leg, the clinician thinks: 'Here's an *acute*, *immediate* lower back pain radiating *below-the-knee* pain, with no *systemic* manifestations; he's more likely to suffer from a herniated disc with a compression than a sprain, an inflammation, associated with *delayed* and *above-the-knee* manifestations.' Furthermore, links between clinical and basic science knowledge are brought to mind through the semantic qualifiers, such as 'immediate pain', related to a nerve compression, compared with 'delayed pain', related to an inflammation.

In a series of observational studies conducted over a decade, using qualitative think-aloud protocols for neurological, gastrointestinal, intensive care and rheumatological conditions, we observed that successful diagnosticians used semantic qualifiers more frequently and in more diversified sets in their discourses than diagnosticians who were less successful. They 'organised the symptoms and signs into coherent systems of relationships of abstract qualities, [and have] broader and deeper representation and understanding of the problems'.[9] (For a summary and educational implications, see Bordage.[9]) Those with diagnostic difficulties or inaccuracies stuck with a factual, literal view of the case, and failed to see the more abstract semantic dimensions of the problem, expressed with semantic qualifiers. This was true for both medical students and experienced doctors; it represents a common characteristic that cuts across levels of experience and taps into the abstract mental scaffolding used to relate bits and pieces of knowledge.

This work on semantic qualifiers provided the opportunity not only to test structural semantic theory in medicine, but also to contribute to further development of the theory by proposing 4 categories of discourses based on 2 organisational dimensions, a semantic dimension, as measured by the number of semantic qualifiers used, and a syntactic dimension, as measured by the extent (length) of the discourses:

1 reduced discourses (limited semantic content and limited discourses);

2 dispersed (limited semantic content and extended discourses);

3 elaborated (semantic richness and extended discourses), and

4 compiled (semantic richness and limited discourses).[10]

The fourth category is similar to the notion of encapsulation described by Schmidt and Boshuizen.[11] Greater diagnostic accuracy and understanding was associated with semantically rich discourses, either elaborated or compiled.

Programmatic research provides depth of understanding over time

Based on think-aloud protocols and the hypothesis that more successful diagnosticians would use more semantic qualifiers to represent a patient's chief complaint, a case-control observational study was conducted.[12] The results showed that successful diagnosticians (case) used over twice as many semantic qualifiers in their characterisation or representation of the chief complaints (means of 3.8 versus 1.6) as unsuccessful diagnosticians and entertained twice as many diagnoses in their differential (means of 2.9 versus 1.5); for example: 'This looks like a *localised*, *acute*, *recurrent*, *large* joint *mono* arthritis that is more likely

associated with *gout* than a *septic arthritis* or any small joint arthritis like *rheumatoid arthritis*.'[13] The role of problem representation, as portrayed here by the use of semantic qualifiers to construct a more abstract mental image of the problem, is an important analytical strategy for solving problems (see, for example, Brenner *et al.*[14] in mathematics education).

In the late 1990s, we tested an education corollary of the structural semantic theory, hypothesising that promoting elaborated semantic representations of chief complaints during case presentations would improve diagnostic accuracy. The results from this experimental study were mixed.[15] Although the students in the experimental group increased their use of semantic qualifiers to describe their cases, this was not associated with improved accuracy. It is likely that when the students initially learned this knowledge, earlier in their education, it was not organised according to semantic networks, and so although they increased their use of semantic qualifiers, they did not have existing relational semantic networks to tap into. The mixed results led to the design of a new, and yet to be conducted, study based on experimental procedures developed by Woods *et al.*[16] to teach basic science concepts related to clinical knowledge; the same procedures will be used here but to teach semantic networks and scaffolding.

Three main lessons can be gleaned from our work on prototypes and semantic networks over the past 3 decades:

1 theory, practice and research methods constantly interact to provide a dynamic interplay that helps move theory and practice forward;

2 programmatic research provides depth of understanding over time, and

3 the human mind and memory are flexible and versatile entities.

First, medical education is essentially a practical field where theories and evidence can inform practice and vice versa. In an American Education Research Association symposium 15 years ago (1992), the late Terry Mast advocated that: 'Theory should drive research; research should drive theory; practice should drive both research and theory… moreover, research and theory should have some influence on practice' (in Bordage and Williams[17]). In our latest test of semantic theory, mixed results have prompted us to go back and test the theory from a different perspective, thus highlighting the dynamic nature of the research enterprise. Norman[18] has argued that theories are dynamic entities set forth to be proven or disproven and to be built upon. This has been and continues to be the case over the years in our work on prototypes and semantic qualifiers. Similarly, a single method is unlikely to provide satisfactory answers to all questions. Competing hypotheses and multiple, converging methods are more productive than single hypotheses and methods. The theory being tested and the questions addressed should determine which methods to use, whether protocol analyses, observational studies or experimental studies – not the other way around.

The human mind and memory are flexible and versatile entities

Second, theory testing, theory building and depth of understanding are facilitated through long-term, programmatic research as opposed to isolated, opportunistic research. A programme of research allows for concepts and variables to be measured and tested systematically. Most research in medical education today consists of isolated studies conducted without any conceptual or theoretical foundation.[19] Theories should not only frame research questions and educational innovations, although that alone would help greatly in medical education, but should also be tested and built upon. As the interactions among theories, methods and practice evolve, the more likely it is that sound educational principles can emerge to better inform fellow researchers, clinical teachers, administrators, policy makers, and licensing and certification agencies, and eventually have some impact on patient care, the ultimate goal of medical education.

Third, in the beginning I had asked how medical knowledge was organised in memory, as if there were a single fixed, crystallised system of organising knowledge in memory. Research has shown that memory is fluid and flexible. Consequently, knowledge and its organisation can be represented in numerous ways, such as prototypes, either as abstractions, as in our work, or as instances, as in Norman and Brook's work on non-analytical reasoning,[20] or as illness scripts, causal networks and semantic structures, to name a few (see Norman[21] for a review). Eva[22] has shown that learning is maximised by promoting both analytical and non-analytical (pattern recognition) structures and processes during clinical education. The most recent results by Ark *et al.*,[23] showed enhanced diagnostic accuracy when learners compared and contrasted categories, thereby fuelling the authors'

speculations about better understanding the diagnosticity of the signs and symptoms, and raise yet another hypothesis about the potential dynamic role that semantic qualifiers, with their dichotomous and contrasting features, might play in facilitating diagnosis.

Any quest to unravel how knowledge is organised in memory has to include consideration of the versatile nature of the human mind in response to situation-specific demands within and across individuals, which brings us back to the concept of case-specificity introduced by Elstein *et al.*[1] in their Inquiry Project of the 1970s (see Norman *et al.*[24] for a more recent iteration). Basic science and clinical instruction will be best served by recognising both the structural nature of knowledge and the corresponding experiential instances in memory, which emerge in learners from repeated exposure to clinical cases and didactic instruction.[25,26] This is akin to Ericsson's deliberate practice with feedback[27] and the mixed practice described by Hatala *et al.*,[28] where deliberate, mixed practice provides opportunities to build a repertoire of instances, concepts and relationships, and the feedback to highlight the accuracy of the decisions and the structural integrity of the instances, concepts and relationships. As more becomes known about the multiple facets and interdigitations of clinical reasoning and knowledge organisation, more attention can focus on instructional strategies that can optimise clinical reasoning, as exemplified by the work of Eva and collaborators, and, if needed, help identify and offer remediation to medical students and house officers who need assistance in becoming expert clinicians.

REFERENCES

1 Elstein SA, Shulman LS, Sprafka SA. *Medical Problem Solving: an Analysis of Clinical Reasoning*. Cambridge, MA: Harvard University Press 1978;276.
2 Mervis CB, Catlin J, Rosch E. Relationship among goodness of example, category norms, and word frequency. *Bull Psychon Soc* 1976;**7**:283.
3 Rosch E, Mervis CB. Family resemblance: studies in the internal structure of categories. *Cognit Psychol* 1975;**7**:573.
4 Rosch E. Principles of categorization. In: *Rosch E, Lloyd BB, eds. Cognition and Categorization*. Potomac, MD: Erlbaum Press 1978;27–48.
5 Bordage G, Zacks R. The structure of medical knowledge in the memories of medical students and general practitioners: categories and prototypes. *Med Educ* 1984;**18**:406–16.
6 Bordage G. The curriculum overloaded and too general. *Med Educ* 1987;**21**:183–8.
7 Cushing H. *The Life of Sir William Osler*, Vol. 1. Oxford: Clarendon Press 1925;460.
8 Osler W. After Twenty-Five Years. *Montreal Medical Journal* 1899; **28**:823.
9 Bordage G. Elaborated knowledge: a key to successful diagnostic thinking. *Acad Med* 1994;**69**:883–5.
10 Bordage G, Lemieux M. Structuralism and medical education: a comparative study of the cognitive strategies of novice physicians. *Semiotic Inquiry* 1986;**6**:143–79.
11 Schmidt H, Boshuizen H. Encapsulation of biomedical knowledge. In: Evans D, Patel V, eds. *Advanced Models of Cognition for Medical Training and Practice*. New York: Springer Verlag 1992;265–82.
12 Chang R, Bordage G, Connell K. The importance of early problem representation during case presentations. *Acad Med* 1998;**73 (Suppl)**:109–11.
13 Bordage G, Connell K, Chang R, Gecht M, Sinacore J. Assessing the semantic content of clinical case presentations: studies of reliability and concurrent validity. *Acad Med*, 1997;**72 (Suppl)**: 37–9.
14 Brenner ME, Mayer RE, Moseley B, Barr T, Duran R, Reed BS, Webb D. Learning by understanding the role of multiple problem representations in learning algebra. *Am Educ Res J* 1997;**34**:663–9.
15 Nendaz M, Bordage G. Promoting diagnostic problem representation. *Med Educ* 2002;**36**: 761–7.
16 Woods NN, Neville AJ, Levinson AJ, Howey EHA, Oczkowski WJ, Norman GR. The value of basic science in clinical diagnosis. *Acad Med* 2006;**81 (Suppl)**:124–7.
17 Bordage G, Williams R. Relationship between theory and practice. *Prof Educ Res Q* 1992;**14**:1–4.
18 Norman GR. Editorial: theory-testing research versus theory-based research. *Adv Health Sci Educ Theory Pract* 2004;**9**:175–8.
19 Cook D, Beckman T, Bordage G. Quality of reporting experimental studies in medical education: a systematic review. *Med Educ* 2007; **41**:737–45.
20 Norman GR, Brooks LR. The non-analytical basis of clinical reasoning. *Adv Health Sci Educ Theory Pract* 1997;**2**:173–84.
21 Norman GR. Research in clinical reasoning: past history and current trends. *Med Educ* 2005;**39**:418–27.
22 Eva K. What every teacher needs to know about clinical reasoning. *Med Educ* 2005;**39**:98–106.
23 Ark TK, Brooks LR, Eva K. The benefits of flexibility: the pedagogical value of instructions to adopt multifaceted diagnostic reasoning strategies. *Med Educ* 2007;**41**:281–7.
24 Norman GR, Bordage G., Page G, Keane D. How specific is case specificity? *Med Educ* 2006;**40**: 618–23.
25 Bowen J. Educational strategies to promote clinical diagnostic reasoning. *N Engl J Med* 2006;**355**:2217–25.
26 Norman G. Building on experience – the development of clinical

reasoning. *N Engl J Med* 2006;**355**:2251–2.
27 Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Acad Med* 2004;**79 (Suppl)**:70–81.
28 Hatala RM, Brooks LR, Norman GR. Practice makes perfect: the critical role of mixed practice in the acquisition of ECG interpretation skills. *Adv Health Sci Educ Theory Pract* 2003;**8**:17–26.

Assessment: do we need to broaden our methodological horizons?

Ayelet Kuper, Scott Reeves, Mathieu Albert & Brian David Hodges

Although medical education is a broad field of research and practice, it has come to be dominated by issues of assessment. Reasons for this emphasis range from the focus on accountability for educational outcomes[1] to the established relationship between assessment and student motivation.[2] Researchers in the domain, especially in North America, have largely focused on methodologies taken from psychometrics and have overlooked the broader social sciences literature devoted to the analysis of social behaviour and social interaction. In this commentary we provide a critique of the ubiquitous use of psychometric methodologies and perspectives and argue that the social sciences offer other rich methodological resources for the study of assessment.

Within medical education research, evaluation is almost always carried out using a set of appraisal tools that are collectively known as psychometrics. We talk about whether a test is valid (whether it measures the thing we want to measure) and whether it is reliable (whether it measures it in a reproducible fashion). Psychometrics has been very successful in evaluating the assessment of many aspects of medical training. It has, for example, allowed the medical education community to systematically evaluate different measures of medical content knowledge,[3] as well as to show that technical skills can be assessed in a reproducible, valid manner.[4]

The social sciences offer other rich methodological resources for the study of assessment

What is rarely made explicit, however, is what the use of psychometric analysis implies about that which is being assessed. More sophisticated psychometricians do stipulate that the latent traits they measure do not really 'exist in any physical or physiological sense'[5] – that they're 'statistical constructs'.[5] However, there is a longstanding implicit reification in the literature of the existence of these underlying internal traits that can be measured over time.[6,7] In either case, it is clear that psychometric tools were initially developed by cognitive psychologists to be valid and useable for phenomena that could at least be conceptualised as stable traits within a single individual. They were designed for the assessment of personality traits such as intelligence, honesty and diligence. Despite issues of test–retest reliability and other methodological hurdles related to positive and negative changes in knowledge over time, they have since been extended for use in the assessment of knowledge and performance.

There is growing understanding that some aspects of medical education are better thought of as social constructs: instead of being considered as expressions of a single individual's abilities, they are conceived of as the products of interactions between two or more individuals or groups

With this caveat, psychometric tools have proven themselves to be very useful for assessing the aspects of medical training, such as content knowledge, that are more easily conceptualised as psychological constructs, as existing individually within each trainee. There is, however, a growing understanding that some aspects of medical education are better thought of as social constructs. That is, instead of being considered as expressions of a single individual's abilities, they are conceived of as the products of interactions between 2 or more individuals or groups.

The Wilson Centre for Research in Education, University of Toronto, Toronto, Canada

Correspondence: Ayelet Kuper, Sunnybrook Health Sciences Centre, 2075 Bayview Avenue, Room HG 08, Toronto, Ontario M4N 3M5, Canada. Tel: 00 1 416 480 4994; Fax: 00 1 416 480 4563; E-mail: ayelet94@post.harvard.edu

doi: 10.1111/j.1365-2923.2007.02945.x