

- conceptualisation, measurement and correlates in full-time clinicians and academic clinicians. In: Caltone MP, ed. *Handbook of Lifelong Learning Developments*. New York, NY: Nova Science Publishers 2010;37–78.
- 5 Wetzell AP, Mazmanian PE, Hojat M, Kreutzer KO, Carrico RJ, Carr C, Veloski J, Rafiq A. Measuring medical students' orientation toward lifelong learning: a psychometric evaluation. *Acad Med* 2010;**85**:S41–S44.
 - 6 General Medical Council. Revalidation – GMP framework for appraisal and assessment. http://www.gmc-uk.org/doctors/revalidation/revalidation_gmp_framework.asp. [Accessed 13 November 2010.]
 - 7 Maisonneuve H, Matillon Y, Negri A, Pallares L, Vigneri R, Young HL. Continuing medical education and professional revalidation in Europe: five case examples. *J Contin Educ Health Prof* 2009;**29** (1): 58–62.
 - 8 Knowles M. *Self-Directed Learning: A Guide for Learners and Teachers*. New York, NY: Association Press 1975:18.
 - 9 Garrison DR. Self-directed learning: toward a comprehensive model. *Adult Educ Q* 1997;**48** (1):18–33.
 - 10 Pilling-Cormick J, Garrison DR. Self-directed and self-regulated learning: conceptual links. *Can J Univ Contin Educ* 2007;**33** (2):13–33.

Towards greater clarity in the role of ambiguity in clinical reasoning

Stephen Lurie

Experts solve problems differently than do beginners. They rapidly identify patterns, prioritise available data and seek out the information that will most efficiently allow them to reach a solution. They can identify common scenarios that can be resolved with few additional data, as opposed to those that require deliberate reasoning from first principles. In general, they have efficient problem-solving strategies that allow them to arrive rapidly at the most likely set of diagnoses.¹

Experts rapidly identify patterns, prioritise data and seek out the information that will most efficiently allow them to reach a solution

The one thing experts generally cannot do, however, is articulate the mental processes by which they accomplish all this.² Thus, theoretical models are needed to explain how experts think. One model proposes that experts, through long experience, develop intuitive 'scripts' that operate autonomously, beneath the level of conscious awareness. Scripts consist of propositional, probabilistic networks of relationships among large numbers of items. In approaching an ambiguous clinical situation, experts are thought to map the available information onto the most appropriate scripts.^{3,4}

As described by Lubarsky *et al.*⁵ in this issue of *Medical Education*, the script concordance test (SCT) has been developed to assess this con-

struct of expert clinical reasoning. An SCT assumes some basic level of factual knowledge, but extends beyond this to assess how the test-taker relates the probabilities of relationships among various pieces of data in hypothetical clinical scenarios. Examinees' responses are compared with those of a reference panel of experts. Members of the reference panel are expected to disagree among themselves on the 'best' response, thereby reflecting the probabilistic nature of clinical decision making. Test-takers are graded based on the degree to which their responses are concordant with those of the reference panel.

Experts are thought to map the available information onto the most appropriate scripts

Lubarsky *et al.*⁵ should be praised for conducting a thorough and rigorous argument-based assessment of the evidence for the

Rochester, New York, USA

Correspondence: Dr Stephen Lurie, Department of Family Medicine, Office of Curriculum and Assessment, University of Rochester, 601 Elmwood Avenue, Box 601, Rochester, New York 14620, USA. Tel: 00 1 585 273 4323; Fax: 00 1 585 273 3290; E-mail: stephen_lurie@urmc.rochester.edu

doi: 10.1111/j.1365-2923.2011.03938.x

validity of the SCT. This is a particularly difficult task, given that the SCT embraces the very attributes that are generally considered the hallmarks of a poorly constructed test: ambiguity and uncertainty. Thus, it is important to identify those aspects of ambiguity that are integral to the theory underlying the SCT, as opposed to those that reflect common construct-irrelevant variability (e.g. response bias, statistical noise, method variance, cueing, etc.). To some extent this problem can be addressed by making further technical refinements in areas such as the scoring method and the optimal size of the reference panel, and by improving the construction of items.⁶⁻⁸ There are at least three theoretical concerns about the role of ambiguity in expert reasoning, however, that may benefit from more precise definition.

Firstly, it remains unclear exactly how and when students develop from novices into experts. This developmental process is probably non-linear, unlike that of simply amassing an increasing fund of knowledge. (Students may actually know more facts about some areas of medicine than do their teachers.) Students struggle with inefficient problem-solving strategies as a part of their evolution towards becoming expert clinicians.⁹ If some of those strategies yield systematically wrong answers (which would nonetheless indicate the progress of an individual from beginner towards novice status), comparison against the responses of an expert reference panel may not reflect this intellectual development.

Students struggle with inefficient problem-solving strategies as a part of their evolution towards becoming expert clinicians

The explication of stage-appropriate tasks in clinical problem solving would thus represent an exciting new direction in theory and practice. What are the developmentally appropriate reasoning tasks for a Year 2 medical student, for instance, and do the 'right' answers for these tasks carry the same degree of ambiguity as those for expert clinician tasks? Given that a strategy of deliberate practice seems to be necessary to attain expertise,¹⁰ perhaps students at earlier stages should be assessed on their approach to practice, rather than asked to demonstrate actual expertise. In general, the tandem development of theory and practice in this area would set clear milestones for clinical performance, and could thus help to guide rational curricular design.

A second source of ambiguity is that experts are expected to disagree on the best possible response; robust disagreement is in fact the hallmark of a good SCT item, as Lubarsky *et al.*⁵ point out. But how much of this represents the inescapably ambiguous nature of clinical problem solving versus common randomness and response bias? Perhaps some of the people chosen for the reference panel are in fact less skilled than their peers. The criterion of being an 'expert' may threaten to devolve into circularity as there is no independent measure of expertise other than nomination to membership on the reference panel. As Lubarsky *et al.*⁵ state, studies are needed to assess the extent to which members of the reference panel consistently differ from their colleagues. Are experts who consistently disagree with their peers expressing a legitimate point of view, or are they merely analogous to items with poor item-total correlations, whose exclusion would actually increase the reliability of the test? And if such false 'experts' were excluded, how much

legitimate disagreement might remain among the members of an expert panel?

The tandem development of theory and practice would set clear milestones for clinical performance and thus help to guide rational curricular design

On the other hand, perhaps a divergence of opinion among experts reflects reasonable alternative strategies. Indeed, at least one study has found that items that provoke higher levels of disagreement among experts are also those that best discriminate among test-takers at different levels of experience.⁶ (Such items, however, also appear to be more difficult, which may independently increase their ability to discriminate experts from novices.) If anything, however, this conjecture makes the problem even more complex because it implies that there are many ways to be an expert. Criteria are needed to discriminate 'true' experts from those whose diagnostic approach may be internally coherent, but also idiosyncratic, inefficient and error-prone. Such individuals would be analogous to a testing strategy that is reliable, but not valid.

Perhaps a divergence of opinion among experts represents an honest alternative strategy for how an expert might solve a problem

A third related source of potentially unwanted ambiguity is that the SCT, with its clearly stated sequential choices and response options, does not mirror the complex way in which information is presented in actual clinical encounters. In a real-time clinical situation, novices may be so overwhelmed by distracting information that they fail to process important data. Thus,

paradoxically, the SCT may present them with a less ambiguous set of choices than they would face in an analogous clinical situation. By contrast, experts may actively direct the clinical encounter in order to seek out the most discriminating data; such behaviours may not be reflected in the predetermined SCT response options. Thus, in an SCT, experts may be faced with an oversimplified scenario which does not necessarily reflect the situation that would have evolved if they had been allowed to direct the clinical encounter. These problems do not 'invalidate' the ability of the SCT to assess clinical reasoning, but they do represent unwanted sources of ambiguity that are unrelated to people's 'actual' degree of clinical reasoning. Future refinements in the mechanics of delivering SCT questions may seek to minimise these sources of unwanted ambiguity.

These considerations lead to basic questions about the nature of clinical expertise. We speak of the 'art' of medicine, an ineffable quality that is reflected in the deliberate use of ambiguity in the SCT. For true artists – painters, poets, musicians – expertise lies in the ability to express a unique and coherent point of view, beginning from the complete ambiguity of a blank page or canvas. As with clinical experts, we distinguish such artists from protocol-driven practitioners. The latter may be competent at their craft, but their work is formulaic, predictable and unoriginal. By contrast, artists who transcend the technical aspects of their craft cannot easily be compared with one another. There is no objective criterion by which one could finally demonstrate that Brahms is a better composer than Bach, or that Klee is a better painter than Kandinsky (although connoisseurs would readily choose the work of any of them over that of

an amateur). Such judgements cannot be settled by reference to some evidence-based standard. They hinge upon particulars of personal history, preference and, not to a small degree, whatever mood one happens to find oneself in at the moment. In the end, there is simply no accounting for taste.

Protocol-driven practitioners may be competent at their craft, but their work is formulaic, predictable and unoriginal

Patients would be justifiably worried to think that expert clinicians were as individualistic in their approach to clinical problems. To some extent, the difference between clinical and artistic expertise represents the fact that there is a much smaller set of defensible approaches to solving a clinical problem than there might be to, say, writing a poem. Moreover, the goal of solving a clinical problem is to limit uncertainty, whereas a work of art may succeed precisely because it creates fertile new doubts.

The critical ambiguity in clinical reasoning is thus related not so much to differences between experts, but, rather, to the infinite variability among clinical presentations with similar underlying etiologies. Although expert clinicians might take somewhat different pathways from data to diagnosis, they should rapidly converge upon a set of similarly ordered differential diagnoses. Thus, it appears that the critical feature in expert clinical reasoning is the efficiency with which the expert's strategy wrests useful coherence from ambiguity. Time, efficiency and accuracy can be easily measured and may represent useful additional criteria for expert reasoning.¹¹

REFERENCES

- 1 Norman G. Research in clinical reasoning: past history and current trends. *Med Educ* 2005;**39** (4):418–27.
- 2 Gladwell M. *Blink: The Power of Thinking without Thinking*. New York, NY: Little, Brown & Co. 2005;277.
- 3 Charlin B, Roy L, Brailovsky C, Goulet F, van der Vleuten C. The script concordance test: a tool to assess the reflective clinician. *Teach Learn Med* 2000;**12** (4):189–95.
- 4 Charlin B, van der Vleuten C. Standardised assessment of reasoning in contexts of uncertainty: the script concordance approach. *Eval Health Prof* 2004;**27** (3):304–19.
- 5 Lubarsky S, Charlin B, Cook DA, Chalk C, van der Vleuten CPM. Script concordance testing: a review of published validity evidence. *Med Educ* 2011;**45**: 329–38.
- 6 Charlin B, Gagnon R, Pelletier J, Coletti M, Abi-Rizk G, Nasr C, Sauv e E, van der Vleuten C. Assessment of clinical reasoning in the context of uncertainty: the effect of variability within the reference panel. *Med Educ* 2006; **40** (9):848–54.
- 7 Gagnon R, Charlin B, Lambert C, Carriere B, van der Vleuten C. Script concordance testing: more cases or more questions? *Adv Health Sci Educ Theory Pract* 2009;**14** (3):367–75.
- 8 Charlin B, Gagnon R, Lubarsky S, Lambert C, Meterissian S, Chalk C, Goudreau J, van der Vleuten C. Assessment in the context of uncertainty using the script concordance test: more meaning for scores. *Teach Learn Med* 2010;**22** (3):180–6.
- 9 Nendaz MR, Bordage G. Promoting diagnostic problem representation. *Med Educ* 2002;**36** (8):760–6.
- 10 Ericsson KA. Deliberate practice and acquisition of expert performance: a general overview. *Acad Emerg Med* 2008;**15** (11):988–94.
- 11 Gagnon R, Charlin B, Roy L, St-Martin M, Sauv e E, Boshuizen HP, van der Vleuten C. The cognitive validity of the Script Concordance Test: a processing time study. *Teach Learn Med* 2006; **18** (1):22–7.