Reviewer's report

Title: Using a conceptual framework during learning attenuates the loss of expert-type knowledge structure.

Version: 1 Date: 30 January 2006

Reviewer: Kevin Eva

Reviewer's report:

General
34 students were assessed via self-report questionnaire regarding whether or not a program-provided diagnostic scheme was used during learning. The 24 that used the diagnostic scheme revealed a higher proportion of “expert-type knowledge structure” upon completing a concept-sorting task than did the 10 who did not learn via the diagnostic scheme. The difference was only technically significant after long-term follow-up, but the effect appears substantial at both points in time. Apparent knowledge structure changed (unpredictably) from expert to novice or vice versa over time in 53% of students.

This is an interesting paper, with publishable results. There are a few things the authors should address, however, in order to improve the work and set it in its proper context.

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)
- Drawing inferences from concept sorting procedures: The claim made on page 6 is that the categorizations individuals generate and the labels they apply indicate underlying knowledge structure. However, psychologists have shown that people can be given a completely arbitrary set of stimuli and successfully generate an appropriate category even if the category has never been considered in the past. Findings like these raise question of whether accurate inferences can indeed be drawn about knowledge structure from the categorizations. The argument in favour of the methodological approach used should be strengthened, perhaps with reference to and description of McGaghie, et al. (2004) and Schmidt (2004), both of which are in Advances in Health Sciences Education.
- Reliability of concept-sorting procedures: It seems clear that raters agree with one another quite well regarding how a particular concept sort should be categorized, but the important reliability question pertaining to this methodology is, how likely is it that someone will give the same sort on two different occasions? If one's sort changes from day-to-day or week-to-week it would be further indication that, in fact, the sorts are not truly indicative of underlying mental structures. Given that the majority of students changed between time 1 and time 2, but that the change wasn't in a systematic direction, one must question whether or not the change is being heavily driven by error of measurement. That is, is one of the predominant “other variables” alluded to in the first paragraph on page 10 simply error attributable to unspecified (and potentially random) forces that influence what concept sort subjects provide on any particular day?

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)
- Methodological details:
  - How do the “concepts” relate to figure 1? Is each of the 13 boxes a concept? In determining whether or not a learners' concept sort is expert-like, does the hierarchy/position of the items matter or is it sufficient to simply name the same concept that experts name? What impact is the decision
likely to have on the results?
o Four concepts were identified as being “expert” concepts - what’s the denominator? That is, does
4 indicate substantial agreement, or the few points that people could agree on in an ocean of
variability?
o Why dichotomize into expert / non-expert based on number of statements? Expertise is very likely
to be a continuum and your analyses should have more statistical power if you compare independen
ties to the continuous measure of number of expert-identified concepts.
- In the regression analyses “short-term knowledge structure” was used as an explanatory variable,
but there is no direct mention of whether or not it was predictive of long-term knowledge structure by
itself. Table 1 would suggest the answer is no. That raises question of the implications of this work -
is it useful to assess knowledge structure if current knowledge structure is unpredictable of future
knowledge structure (perhaps especially given the authors claim that knowledge structure is
predictive of problem-solving success). Can “knowledge structure” as it is measured here truly be
considered a proxy measure of ability/expertise if it's non-stable?
- The interaction between short-term knowledge structure and use of a diagnostic scheme is
described, but not in a way that relates to the data presented. The authors claim that “students who
develop expert-type knowledge structure during the period of instruction and who use a diagnostic
scheme in doing so were more likely to keep expert-type knowledge” (p.10/11, underline added). In
fact, however, the data illustrated in table 2 does not show the proportion of students who showed
expert-type knowledge in the long-term as a function of the expert-type knowledge they held in the
short-term. Table 2 should be a replication of table 1, but for only the 24 who used diagnostic
schemes, and then again for only the 10 who did not use diagnostic schemes. It’s important to the
claims made to know these proportions given that table 1 makes it clear the 58.3% in table 2 are not
entirely drawn from the 63.8% above it. This relates to the implications differing when one talks of
“change in expert-type reasoning” relative to a “decline in proportion of the number of students using
expert-type reasoning” that is described below.

Discretionary Revisions (which the author can choose to ignore)
- The paper would benefit from a stronger claim to the relation between expert sorts and aptitude in
this domain. Are the renal course performance scores available? Clerkship scores?
- An important wording issue: Throughout the paper the authors refer to “change in knowledge
structures.” On page 9, for example, and the corresponding discussion, the authors state that the
group who did not use a diagnostic scheme showed a “significant change in knowledge structure
towards novice type knowledge” - this implies that individuals showed a change towards that type on
average yet the authors acknowledge that many individuals changed in the opposite direction. More
appropriate would be to discuss these effects as “there was a significant decrease in the proportion
of students using expert-type knowledge structures” throughout the manuscript.
- The opening sentence is certainly the ideal, but in fact, one of Needham and Begg's points is that
the transfer described doesn't occur 100% of the time as is implied by this sentence
- Claiming that “The” diagnostic scheme is shown in Figure 1 demands a description of the actual
reliability of experts’ judgments in this regard and an indication of how this particular scheme was
derived.
- Typos: Page 5, line 1 - “their” should be “there”; Page 11, 4th last line - “a” should be “as”; Page 18
(Table 2) - it’s impossible for 10 people to be divided into to make 40.9% - I assume that should be
40.0.
- Please include the actual proportions along with the percentages throughout the paper to better
enable readers to assess the results: On page 8, 44% and 57% should, for example, be listed as
15/34 and 19/34 in addition to the percentages. This change should be made throughout the text as
well as in the tables.
What next?: Accept after minor essential revisions

Level of interest: An article of limited interest

Quality of written English: Acceptable

Statistical review: No

Declaration of competing interests:

I declare that I have no competing interests