Réponses à l'éditeur concernant l'article de Bland

Bernard Charlin

We read with interest Charlin's response to our recent article in Academic Medicine (1) regarding the psychometric properties of the script concordance test for Nephrology. Charlin offers another hypothesis for the possible advantage of a 3-point Likert-type scale over a 5-point scale: The problems we asked the examinees to solve were not ill defined.

During the design process, we included no pathognomonic signs, symptoms or studies in the data. We constructed the cases based on actual patients who presented a diagnostic challenge to the question writer during his or her practice of nephrology and our expert reviewers flagged any items that were not felt to offer a diagnostic challenge to an expert.

The sixteen experts agreed unanimously on only one of our fifty questions. Using a five point scale the mean number of experts choosing the modal answer was 8.28 with 75% of questions having between 10 and 6.75 experts selecting the mode as the answer with a standard deviation of 2.65.

Utilizing the re-coded 3-point scale, 10.74 experts chose the modal answer with 75% of the questions having between 9 and 13 experts selecting the mode as their correct answer with a standard deviation of 2.62.

Clearly, the expert panel unanimously felt a correct answer existed for only one of fifty items. We feel this data along with our construction method created ill-defined problems for our examinees to solve.

If Charlin were to re code and reanalyze his SCT data, would we discover a similar trend? Charlin's answer keys have evolved from a 7-point Likert- type scale to a 5-point Likert-type scale. Our data suggests there is little information lost in going from a 5-point scale to a 3-point scale. The 3- point scale does not require the examinee to be certain the new information proves or disproves the hypothesis. Rather, we ask the examinee if it increases or decreases the probability of an outcome.

It is possible that the strength of the SCT lies in its question construction rather than its Likert-like aggregate scoring algorithms. The micro- decisions required to complete the SCT may detect the acquisition of clinical problem solving skills earlier than MCQ's.

In a recent abstract presented at the CGEA, the SCT was administered as a Pre/Posttest to 15 residents immediately before and after a nephrology rotation. Another group of 15 residents were tested one-year after the original SCT administration. (2)

The Pre/Post scores demonstrated no change in performance(- 0.2%p=0.9). However, the 1-year interval group showed a statistically significant increase in scores(+10% p=0.021).

The significant change in the 1-year interval group and not the Pre/Posttest group suggests SCT measures the development of scripts used in clinical problem solving rather than rote knowledge. Knowledge test scores would be expected to increase immediately after the rotation, but the elaborated, interconnected network of knowledge and experience as measured by the SCT is predicted by theory and observation to take longer to develop.

These findings suggest the SCT may be a measure of clinical reasoning rather than rote knowledge. We agree with Charlin's call for further research on this new and potentially powerful assessment tool.

References

1. Bland AC, Kreiter C, Gordon JA: Investigating the Psychometric Properties of Five Scoring Methods Applied to the Script Concordance Test. Academic Medicine 2005;80:395-399

2. Sader S, Bland AC, Cation L, Kreiter C, Hendricker R, Antes L, Aiyer M, Gordon JA. The Script Concordance Test for Nephrology: A Tool to Assess Experience in Clinical Problem Solving. Central Group on Educational Affairs Spring Meeting 4/2005.