

Judgement biases in assessment: for better or worse? (0286)

Lisa Schmidt, Lambert Schuwirth, Svetlana King
Flinders University, Australia

Judgement is a necessary component of assessment, both to determine the standard of performance and to inform feedback. There have been attempts to avoid subjective judgement via so-called 'objective' methods but this can be at the risk of validity, especially outside the realm of factual knowledge (Schuwirth & van der Vleuten, 2006). In particular, considering Miller's pyramid (knows, knows how, shows how, does) (Norcini, 2003), in the area of assessing the ability of a student to perform an action, it is expert judgement by experienced practitioners that is required.

Human judgement is the method for ensuring valid assessment in certain areas but it is subject to judgement biases. Biases are not prejudices, instead, they are misrepresentations in the assessor's mind of what occurred during the assessment exercise. Some examples include memory errors (by omission or commission), the primacy effect (the impact of the first impression), the recency effect (the impact of the final impression), the salience effect (the impact of the most noticeable or observable thing), selective perception (preferential attention to the 'pet' thing of the examiner), the halo effect (transfer of one attribute of a person to another attribute) and the contrast effect (the impression of one student impacting on the impression of another student) (Plous, 1993). Any of these might impact on an assessor's judgement of a student and shift grades up or down. The potential impact on grades implies that biases are bad and we should try to remove them, but should we?

These biases are hard-wired and have served humans very well – we've survived as a species for this long with them. This implies that they are not necessarily entirely bad things. Consider the primacy effect. A student may walk into an oral examination casually dressed and displaying some social awkwardness. This may give the examiner a negative impression of the student but should this impact on the judgement or not? The answer depends on the context. If it were a science examination on critical analysis of the scientific literature, then appearance and social skills are irrelevant to the purpose of the examination and it would be invalid and inappropriate to factor in dress and social skills. However, if it were a law examination on client interviewing, then the first impression should be taken into account and should alert the examiner to potential concerns around professionalism. So biases may distract assessors or alert them. The underlying issue is that assessors are making decision in real-world environments rather than the highly structured, controlled environment of the bias-discovering psychology research laboratory (Klein, 2008).

This begs the question, "Should we train assessors about biases?" We have taken the view that we should and we outline our training package below. However, we also take the view that we are not even trying to 'train-out' biases as this would be futile. Instead, we aim to raise assessor's awareness of biases so that they can recognise when they might be in a situation where a bias may occur. The awareness then enables them to consciously decide what to do with the information they have. This in turn enables them to justify and articulate the decision they have made and – as illustrated by the legal profession – it is the articulate justification for a decision that provides the reliability. Our analysis of the workshops we have run so far indicates that the language that people acquire through the training is what is empowering – both in terms of their judgement and in terms of teaching teams being able to discuss assessment.

Our training package is called 'Better Judgement' and has been funded by the Australian Government Office for Learning and Teaching. The pedagogy used in the Better Judgement training program is based on current insights from cognitive psychology on effective human

learning. A well-organised knowledge base is essential for the ability to solve problems or deal with complicated situations (Chase & Simon, 1973; Posner, 1988). The training program therefore includes video presentations to explain judgement biases and how they might influence assessors. Also, in expertise development the element of transfer plays an important role (Eva, 2003; Eva, Neville, & Norman, 1998; Posner, 1988); where transfer is the ability to recognise similarity between two different but related problems. The training programme therefore includes sufficient variations of cases in which a certain bias occurs. For transfer to be maximal not only variation of obvious cases must be used but also subtle cases. Participants practice identifying biases in action in well-defined situations using videos of actors in cartoonishly obvious examination scenarios. Participants can then practice in less defined situations using real-life clips found on YouTube. Furthermore, active involvement with tasks leads to better storage, retention and retrieval of what is learned (Freeman et al., 2014). This is mainly included via participants writing scenarios located in their own teaching setting. The training package is available at [*withheld for blind submission review*].

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4(1), 55-81.

Eva, K. W. (2003). On the generality of specificity. *Medical Education*, 37(7), 587-588.

Eva, K. W., Neville, A. J., & Norman, G. R. (1998). Exploring the Etiology of Content Specificity: Factors Influencing Analogic Transfer and Problem Solving. *Academic Medicine*, 73(10), S1-5.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., et al. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*.

Klein, G. (2008). Naturalistic Decision Making. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50(3), 456-460.

Norcini, J. J. (2003). Work based assessment. *BMJ : British Medical Journal*, 326(7392), 753-755.

Plous, S. (1993). *The psychology of judgment and decision making / Scott Plous*. New York: McGraw-Hill.

Posner, M. (1988). What is it to be an expert? In M. Chi, R. Glaser & M. Farr (Eds.), *The nature of expertise* (pp. xxix - xxxvi). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Schuwirth, L. W. T., & van der Vleuten, C. P. M. (2006). A plea for new psychometric models in educational assessment. *Medical Education*, 40(4), 296-300.