



Oxford University Centre for Educational Assessment

THE APPLICATION OF TVERSKY AND KAHNEMAN'S THEORY OF HEURISTICS AND BIASES TO EXAMINERS' DECISION MAKING

VICTORIA ELLIOTT, SUPERVISED BY PROFESSOR I. LUNT & PROFESSOR G. STANLEY

Kahneman and Tversky's theory of heuristics and biases provides an explanation of human ability to make intuitive judgements when the cognitive load is beyond their rational capability. Heuristics are 'rules of thumb' that essentially rely on the substitution of one attribute for another, more difficult attribute, which would be the logical cue for judgement. 'Bias' occurs when cues for judgement are wrongly weighted in the judging process (Frederick & Kahneman, 2002). The examples given below tend to be those where the heuristics lead to corresponding biases, since these illustrate the case more clearly, and it is in these situations where it is easiest to identify heuristics at work. Despite this prevalence, heuristics are a valid and well-used, automatic strategy in human judgement processes.

Affect, representativeness and availability all rely on 'substitution' of one attribute for another attribute, more difficult to assess, which would be the rational cue for judgement. *Anchoring & adjustment*, on the other hand, relies on comparison of the case in hand to a previous secure judgement, or fixed point of reference. The choice of reference point alters the final judgement.

Heuristics are not a conscious strategy for decision making, so that examples given below which appear to be malpractice may be happening without the examiner's knowledge. Indeed, in some cases, if the examiner were aware of the thought prompting their judgement, they would mediate it in order to avoid bias.

"This candidate has terrible handwriting. And he can't spell at all."

REPRESENTATIVENESS Items which appear similar are assumed to have the same characteristics, and an item which appears to fit into a group is assumed to have the characteristics of that group (theorised by Tversky & Kahneman, 1974).
Application: A script whose content is worthy of an A grade may be under-judged because the handwriting and spelling are considered to be sub-standard. Equally a well-presented script may appear representative of a higher grade than its contents require.

"What does full marks look like? This is very good, but is there better?"

AVAILABILITY The ease with which an example can be brought to mind alters the prediction of the frequency of an event, or of the proportion of the population (theorised by Tversky & Kahneman, 1974).
Application: An examiner marks an essay which should receive full marks. However, she might predict that very few or no essays should do so, having never seen an example, and give it less than it deserves.

"Well her first essay was only an E. I don't have high hopes for the second."

ANCHORING & ADJUSTMENT the tendency to form initial judgements around a base point or 'anchor' and then adjust according to the situation and developing information, to reach a final decision. Using an inappropriate anchor and/or failing to adjust sufficiently leads to bias (theorised by Tversky & Kahneman, 1974).

Application: The process by which an examiner sites an answer within a band (anchor) and then fine-tunes the mark up or down (adjustment) for the final judgement. Anchoring effects include using the first of two essays in a script as an anchor for the second, for example, or using the initial paragraph as an anchor.

"What a ridiculous way to interpret this poem! She can't even quote properly!"

AFFECT This heuristic suggests that an emotional response, perhaps unconscious, by the examiner to the quality of an examination script, could dominate the judgement which is made, which is then justified by reasoned argument (theorised by Finucane *et al.* 2000).

Application: An examiner likes or dislikes a candidate's phrasing, handwriting, use of quotation, or virtually any other aspect of their writing, and this affective response dominates the judgement which is made. This is more damaging if the affective response is not consciously recognised, so that compensation is made.

ABSTRACT

This poster demonstrates part of the theoretical model for my D.Phil. project 'The decision making processes of examiners' (supervised by Professor Ingrid Lunt and Professor Gordon Stanley). Examining is supposed to be a rule-based, highly structured and highly rational activity, supporting the requirements of fair assessment. The project considers essay based examinations, such as those for History and English GCE, where the cognitive load required to hold in mind a lengthy essay and a generic one or two page rubric would seem to exceed what may be reasonably used and processed, if Simon's principle of 'bounded rationality' (1992) is invoked. Data will be gathered from examiners while they make decisions using 'think aloud' protocols, and by analysing annotations made on examination scripts.

REFERENCES

- Finucane, M.L., Alhakami, A., Slovic, P. & Johnson, S.M. (2000) 'The affect heuristic in judgements of risks and benefits', *Journal of Behavioral Decision Making*, 13, pp. 1-17
- Kahneman, D. & Frederick, S. (2002) 'Representativeness revisited: attribute substitution in intuitive judgement', in Gilovich, T., Griffin, D. & Kahneman, D. *Heuristics and Biases: the Psychology of Intuitive Judgement* (Cambridge: CUP), pp. 49-81
- Simon, H. (1992) *Economics, Bounded Rationality and the Cognitive Revolution* (Aldershot: Edward Elgar Publishing)
- Tversky, A. & Kahneman, D. (1974) 'Judgement under uncertainty: heuristics and biases', *Science*, 185, pp. 1124-1131

