





OXFORD UNIVERSITY CENTRE FOR EDUCATIONAL ASSESSMENT


Cognitive cherry picking: the patchwork process of examining A level essays



v.f.elliott@warwick.ac.uk
 Cambridge Assessment 6.9.12

THE UNIVERSITY OF **WARWICK**



OXFORD




WARWICK


OXFORD


Methods/ Design

Two subjects: English and History, during the January examination session.


Recording of: one pre-standardisation meeting; four standardisation meetings (two live, two 'voice over internet protocol'; two large, two small).

Three sets of three hours of 'think aloud' data during live marking (Lupin, Parrot and Casper).

WARWICK

UNIVERSITY OF OXFORD

Institute of EDUCATION



Heuristics and biases

WARWICK

UNIVERSITY OF OXFORD

Institute of EDUCATION

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.

WARWICK

UNIVERSITY OF OXFORD

Institute of EDUCATION

"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.

WARWICK

UNIVERSITY OF OXFORD

Institute of EDUCATION

"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. (theorised by Tversky & Kahneman, 1974).

Application: 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.

WARWICK

UNIVERSITY OF OXFORD

Representativeness/ availability

Institute of EDUCATION

you've got several scripts now that are in level 5, and you might go back and think well, this is a far far better level 5 than the William the Conqueror one. That doesn't mean that the William the Conqueror one is *not* level 5. You see, it does display the necessary qualities to get into that level. The fact that [the other] two scripts demonstrate those qualities in spades is unimportant. (Principal Examiner, History 1)

WARWICK

UNIVERSITY OF OXFORD

representativeness

Institute of EDUCATION

Principal Examiner: can we look at J? I found this a bit of a problem one - it's a very short answer

Examiner: Oh golly
[they read]

Examiner: Good stuff though (.) hard to take exception to any of it. Hardly a word wasted. I can't believe that wasn't planned. They probably spent more time thinking about it than writing

Principal Examiner: But it's short isn't it

Examiner *: It is but it's small writing

Examiner: It's literally getting twice the number of words to the line

Senior Examiner: It's discriminating, no doubt, not wasting a word

WARWICK

UNIVERSITY OF OXFORD

Institute of EDUCATION

"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 Fuscane *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.

WARWICK

UNIVERSITY OF OXFORD

Institute of EDUCATION



comparison


WARWICK

UNIVERSITY OF OXFORD


Institute of EDUCATION

that's putting it above M. Not as far as L [9 second pause] but looking back over M... it's better than M – no it is better than M. What did M get? Thirty five. So no, that's not over marking it.
(Lupin, think aloud)

WARWICK

 UNIVERSITY OF OXFORD

comparison



Something else to consider is that this time you're marking both of the items together. If you've awarded marks for both answers which are substantially different from each other it is worth reflecting on your decisions before moving on. In the summer in general, it became obvious that most candidates tend to approach both answers at the same level of um answer and therefore if you've, if you do have an answer which um has a level five or a level three for example then it's worth reflecting on your decisions.

WARWICK


 UNIVERSITY OF OXFORD

Comparison (imagined script)




I think the weakness of both of the ones that we've looked at so far was something that you know, really top candidates will have actually seen in that question which wasn't addressed by either of them. Would you sort of think that there's anything that both of them missed out on that you might be expecting from a really strong candidate?

WARWICK

 UNIVERSITY OF OXFORD

Imagined script subordinated to real script




obviously I think we need to keep in touch on that always though, because it may well be that virtually every candidate, er, interprets the question in that way and er, er, we don't want to debar people from getting a really high mark, if everyone's interpreting it that way. I think we need to keep in touch, on the way people actually answer that question.

WARWICK

UNIVERSITY OF OXFORD

Institute of EDUCATION



Mental frameworks

WARWICK

UNIVERSITY OF OXFORD

Mental frameworks

Institute of EDUCATION


I've usually not got a firm figure in my mind the first time I read through... more a region... particularly not at this stage (.) I mean 50 scripts down the line (.) you know (.) it's more intuitive Unless you try it (.) you can't do it didactically (.) you've got to [makes motion with pen indicating several lines]

As you hit about 50 you've got it sussed

WARWICK

UNIVERSITY OF OXFORD

Institute of EDUCATION



WARWICK



UNIVERSITY OF
OXFORD



"I believe it was Plato who said that good judgement consists equally in seeing the differences between things that are similar and the similarities between things that are different."
Brian Magee, Confessions of a Philosopher

My thanks to Gordon Stanley and Ingrid Lunt for their advice, the ESRC for funding and huge thanks to all who participated in my research.

WARWICK
