



**Cambridge
Assessment**

The effect of adaptivity on the reliability coefficient in comparative judgement

Conference Paper abstract

Sylvia Vitello & Tom Bramley

Presented at the 18th annual AEA-Europe conference,
Prague, Czech Republic,
November 2017

Author contact details:

Sylvia Vitello
Assessment Research and Development,
Research Division
Cambridge Assessment
1 Regent Street
Cambridge
CB2 1GG
UK

vitello.s@cambridgeassessment.org.uk

<http://www.cambridgeassessment.org.uk>

As a department of Cambridge University, Cambridge Assessment is respected and trusted worldwide, managing three world-class examination boards, and maintaining the highest standards in educational assessment and learning. We are a not-for-profit organisation.

How to cite this publication:

Vitello, S. and Bramley, T. (2017, November). *The effect of adaptivity on the reliability coefficient in comparative judgement*. Paper presented at the 18th annual AEA-Europe conference, Prague, Czech Republic.

Abstract

Comparative Judgement (CJ) is an increasingly widely used method for creating a scale, for example of the quality of essays. One popular approach for optimising the selection of pairs of objects for judgement is known as Adaptive Comparative Judgement (ACJ). It has been repeatedly claimed in the literature that ACJ produces very high reliability, often higher than can be obtained by conventional marking. It has been shown by simulation that adaptivity can substantially inflate the apparent reliability in ACJ. The aim of this study was to see if the same inflation would happen in real data by comparing an adaptive with a non-adaptive CJ study using English essays. An all-play-all set of comparisons of a subset of the essays allowed the extent of scale inflation to be quantified: the reported ACJ reliability was 0.97 whereas the all-play-all value was 0.82. The value from the non-adaptive study was 0.70. However, the scale from the non-adaptive study correlated slightly higher with external variables (the Principal Examiner's mark for the essay that was judged, and for a different essay), suggesting the non-adaptive study was no less valid than the adaptive one. The advantages and disadvantages of using adaptivity in CJ will be discussed.