



CAMBRIDGE ASSESSMENT

How well can teachers assess pre-university students' research reports? Findings from an empirical study

A paper presented at the European Conference for Educational Research, Helsinki, Finland, August 2010.

Irenka Suto and Stuart Shaw

E-mail: suto.i@cambridgeassessment.org.uk

Abstract

General description: In the UK and internationally, many students preparing for university are given the challenge of conducting independent research and writing up a report of around 4000 or 5000 words. Such research activities provide students with opportunities to investigate a specialist area of study in greater depth, to cross boundaries with an inter-disciplinary enquiry, or to explore a novel non-school subject such as archaeology, cosmology or anthropology. We theorise that, as is the case in higher education (Brown et al. 1997), independent research encourages intellectual curiosity whilst enabling students to develop skills in practical and analytical research, higher order thinking, interpretation and time management. When applying to university, students can use their reports to demonstrate motivation for their intended course of study and to differentiate themselves from competing applicants.

In the wake of the recommendations of the Tomlinson Report (2004) on the shape of 16-19 qualifications in England, The Sixth Form College, Farnborough, developed a systematic approach to encouraging its students to conduct independent research. The students carry out extended projects during their holidays, or alongside their other courses. They generate formally-structured reports which are assessed formatively through detailed written comments to the students by their teachers. This has generated a considerable body of student evidence within the college. At other schools, students conduct projects which constitute or contribute to a formal qualification, and which are therefore assessed summatively. For some of these qualifications, the students' research reports are assessed by their own teachers. The teachers' marks are then moderated by professional examiners who are employed by the examination board administering the qualification. Examples of this assessment approach include the Cambridge Pre-U Independent Research Report, administered by Cambridge International Education, and the extended projects administered by the AQA, OCR, and Edexcel examination boards. These extended projects can be used to obtain a stand-



UNIVERSITY of CAMBRIDGE
Local Examinations Syndicate

alone qualification. Alternatively they can contribute to a 14-19 Diploma in England or the Welsh Baccalaureate qualification in Wales. For other qualifications, such as the International Baccalaureate, students' research is marked exclusively by external examiners.

The assessment of research reports poses several challenges, including the multi-faceted contradictions that arise when assessment schemes are designed to reward generic research skills rather than particular subject knowledge. Assessors may lack detailed understanding or marking experience of the research topics explored by some students. However, it is unclear whether subject knowledge facilitates or hinders marking. For example, familiarity with particular terminology or technical language may aid interpretation of what the student has written. Alternatively it may obscure the assessor's perception of generic skills, especially if they have been mis-applied by the student.

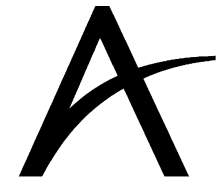
In this study, we sought to investigate these and related concerns, including the feasibility of applying a single mark scheme to research reports covering diverse topics in order to reward generic research skills. Our aim was to investigate the reliability with which teachers can mark diverse research reports, using four different generic assessment objectives. We also investigated teachers' views in applying generic mark schemes, particularly when marking reports on unfamiliar topics.

Methods: The study involved fifteen teachers with diverse subject specialisms from a range of secondary schools across England. All of the teachers were currently or soon to be supporting 16-19 year old students to carry out independent research. Additionally, an experienced professional examiner participated as the Chief Examiner in the study.

Each teacher received an identical sample of twenty 5000-word research reports covering a wide range of topics, and a mark scheme designed to reward generic skills in analysis, evaluation, and communication, and uniquely, intellectual challenge. They were asked to: (i) mark the first five reports; (ii) receive telephone feedback on their marking from the Chief Examiner; (iii) mark the remaining fifteen reports; and (iv) complete a questionnaire about their experiences.

Outcomes: All fifteen teachers marked all twenty reports in the study. Indices of inter-rater reliability among all teachers were calculated. They were found to be encouraging, comparing favourably with other reliability indices for marking extended prose (Johnson, Nádas and Bell, 2009; Laming 1990; Shaw 2008). A multi-faceted Rasch analysis indicated a generally well-fitting Rasch model with no mis-fitting teachers, suggesting similar individual variability among all fifteen teachers. When considered together with the teachers' mean marks and the estimations of inter-rater reliability, the Rasch findings revealed a good degree of agreement among teachers on each of the four generic skills.

After completing their marking, fourteen out of fifteen teachers filled in a questionnaire about their experiences. The questionnaire provided insights into the reasons why teachers perceived some reports to be more difficult/easy to mark than others as well as helping to identify issues relating to marking generic skills.



CAMBRIDGE ASSESSMENT

We report on the study's findings, focusing on teachers' perceptions of the marking process. We discuss the findings' implications for the design of assessments of student research projects, and their relevance to other European countries with similar assessment challenges.

Keywords: extended project, reliability, research project, marking, teacher assessment

References:

Brown, G., Bull, J. and Pendlebury, M. (1997) *Assessing student learning in higher education*. Routledge: London and New York.

Johnson, M., Nádas, R. and Bell, J. F. (2009) Marking essays on screen: An investigation into the reliability of marking extended subjective texts. *British Journal of Educational Technology*.