Carry on examining: Further investigation

Kate Tremain

Summary

This paper follows on from work by Tremain (2011) and Meadows (2004) on examiner satisfaction and looks at the difference in intentions to continue examining between examiners with experiences of different marking media and modes of delivery of standardisation. A postal questionnaire was distributed and a response rate of 40.32% was achieved (n = 1210). The questionnaire consisted of three sections: section 1 contained 14 demographic items, section 2 contained 35 items concerning experiences of examining, and section 3 consisted of 25 items related to examiner personality. Analyses suggested that there was no significant difference in intentions to continue between examiners with different experiences of standardising (online, face-to-face, or both). For marking media, too few respondents had marked online only for this group to be included in the analysis, but a significant difference in intentions to continue was found between examiners who had marked using both media and those who had only marked on paper only, with those who had marked on paper only having lower intentions to continue. However, the effect size was small and the difference in the medians was only 0.33, suggesting the difference has limited practical utility. Moreover, the results of this study must be interpreted with caution due to methodological limitations, such as the potential confounds of subject and length of experience, which make it difficult to understand the cause of such a difference, and difficulties accurately classifying standardisation experiences.

Background

Examiners play a vital role in ensuring that candidates receive appropriate grades for their work and AQA is reliant on the willingness of examiners to perform this role. Thus an understanding of the factors influencing examiner job satisfaction is valuable in efforts to maintain or improve examiner retention and recruitment. However, examiner job satisfaction has been the focus of very little research. One of the few studies on examiner satisfaction was conducted by Meadows (2004). The study used a questionnaire survey to investigate attitudes towards and intentions to continue examining and found that, while they wanted greater financial remuneration, most examiners were satisfied with their experiences. However, the role of examiner has undergone substantial change since 2004, for example, with the introduction of online marking and online standardisation. As such, Tremain (2011) conducted a similar postal questionnaire survey, with the aim of updating Meadows's research. When analysing the responses, it became apparent that many examiners were dissatisfied with the changes to marking and standardising procedures - a finding which corroborated earlier work by Chamberlain (2008). However, whether this dissatisfaction with online marking and standardising translates into lower intentions to continue examining has yet to be empirically explored; this paper aims to conduct such an investigation.

Method

A questionnaire consisting of three sections was posted to a sample of 3,001 examiners who had marked for AQA in the summer of 2010. A response rate of 40.32% was achieved (n = 1210).

Section 1 c onsisted of 14 ques tions relating to examining experience: whether they marked online or on paper, whether they were standardised online, their level of seniority and so on (see Appendix 1). Section 2 (see Tremain, 2011) consisted of 35 statements - 15 positive, 17 negative - about examining, including items on pressure, pay, and standardisation, among others. Participants rated their level of agreement with each statement using a four-point Likert scale (strongly disagree, disagree, agree, strongly agree, and not applicable). The mean of participants' responses to three items was used to calculate an "intention to continue examining" scale: I expect to examine for the AQA for the foreseeable future; I intend to examine for the AQA for the foreseeable future; I want to continue examining for the AQA for the foreseeable future. There were also three open-ended questions. Section 3 consisted of 25 items aimed at assessing respondents' personality traits. However, this approach to assessing personality was unsuccessful and this section will not be discussed further.

As it was originally intended to look at the relation between examiner performance and satisfaction, 16 performance groups were created by cross-tabulating mean mark adjustments and mean grade categories. From the 8,273 examiners for whom these data could be collected, a sample of 3,001 examiners was drawn randomly, while maintaining the proportion of examiners in each group. However, the predominance of examiners in group 1 m eant that performance could not usefully be included in statistical analysis. This method also resulted in examiners using CMI+ being under-represented, as they do not receive mark adjustments. For more details of the sample demographics, please refer to Tremain (2011).

Results

For experiences of marking using different media, examiners were divided into three groups: those who had only ever marked online (n = 17), those who had only marked on pa per (n = 660), and those who had used both media (n = 413). As the group of those who had marked online only was so small, particularly in comparison to the other groups, it was excluded from further analysis.

A Mann-Whitney U test was conducted to investigate whether there was a significant difference in intentions to continue examining between examiners who had marked on paper only compared to those who had marked using both media. The results indicated that there was a significant difference *(U* = 125940.00, *p* = .025, *r* = -0.07), with those marking on paper having lower intentions to continue examining (a median score of 3.67) compared to those who had marked using both media (a median score of 4.00). However, with an *r* of .07, this represents a very small effect.

Kruskal-Wallis tests were also used to investigate the relationship between experiences of different modes of delivering standardisation and intentions to continue examining. Examiners were divided into three groups: those who received only face-to-face standardising (n = 230) in the last series for which they performed marking, those who had been standardised online for some of the components they marked, but not for others (n = 283), and those who were standardised entirely online (n = 586). No significant difference was found between the three groups (*H* (2) = 4.39, p = .11).

Conclusions and caveats

This paper sought to explore the impact of different experiences of marking and standardising on examiners' intentions to continue examining. The results indicate that there was a very small, but significant, difference in intentions to continue examining between examiners who marked on paper only compared to those who used both media, with examiners who marked using both media having higher intentions to continue than their paper only counterparts. There was no

significant difference in intentions to continue between examiners with experiences of standardising with different modes of delivery. However, as examiners were only asked how they were standardised in the most recent series for which they had marked, some examiners in the "online only" group may have had previous experience of face-to-face standardising. As such, this finding is not as robust as that for electronic marking. Nevertheless, some difference between the face-to-face only and the two other groups would have been expected, given anecdotal reports.

Both findings – the lack of difference in intentions for standardising, and the higher intentions to continue of examiners who mark using both media compared to paper only markers – are surprising given the amount of dissatisfaction expressed by examiners in Tremain (2011) regarding online standardising and marking. However, a number of methodological issues merit consideration when interpreting these results. First, as online marking and standardisation are recent developments, it is possible that those who have been standardised online only are newer examiners. It seems reasonable that over time, examiners with less intention to continue will drop out, leaving a population of experienced, committed examiners. As such, there is a potential confound in the results. This confound may also apply to the findings for marking media, with examiners who have used both media having marked for more series on average (a mean of 14.79 series) than those who have only marked on paper (a mean of 12.02 series).

It is also worth noting that those subjects that have not yet transferred to being marked online tend to be subjects which demand extended responses, for which the marking requires more judgement and is therefore more challenging. Examiners from these subjects are likely to predominate in the paper only group, whereas a wider range of subjects may be included in the group who have used both media. Consequently, the marking experience is expected to be very different for the paper only group compared to those who have marked using both media, potentially confounding the results. It would also have been useful to conduct several questionnaires with the same examiners over time, measuring intentions to continue as the changes were introduced, rather than trying to compare different groups of examiners. In addition, it is important to recognise that examiners are nested by component, and therefore the data have a hierarchical structure. The potential effects of component have not been investigated here, but it does not seem unreasonable to suggest that intentions to continue may vary by component, regardless of marking medium. Future research would ideally account for the hierarchical nature of the data in the analysis.

An additional limitation is that those examiners with lower intentions to continue due to the changes in procedure may have left prior to the study. It is also worth considering the possibility of a response bias, whereby those examiners who are more committed to their role may be more inclined to respond to the questionnaire. Finally, the sampling method led to the under-representation of CMI+ examiners, which is likely to have contributed to the small sample size for examiners who had only marked online. Resultantly, this group could not be included in the analysis and thus the generalisability of the findings is limited in this respect.

In sum, intentions to continue examining appear to be unrelated to different experiences of standardising, but do appear to be related to different experiences of marking, with examiners who have only marked on paper having lower intentions to continue than those who have marked both online and on paper. Given the amount of dissatisfaction expressed by examiners about online marking and standardising, these findings are surprising. However, the size of the effect of marking medium on intentions to continue is very small, suggesting the finding has limited practical utility. In addition, this study had a number of methodological limitations. Accordingly, these results must be interpreted with caution.

References

Chamberlain, S. (2008). *Examiners' perspectives of online standardisation.* AQA Research Report, RPA_08_SC_RP_046.

Meadows, M. (2004). *Measures of examiner job satisfaction as predictors of intention to continue examining.* AQA Research Report, RC256.

Tremain, K. (2011). *Carry on examining: what predicts examiners' intentions to continue examining?* AQA Research Report, CERP_11_KMT_RP_043.

Appendix 1 – Examiner satisfaction questionnaire: Section 1

Section 1. About Yourself and Your Examining Experience

a. In summer 2011, did you mark electronically or on paper?

Electronically o   On paper o   Both o

b. If you marked only electronically this series, have you marked on paper in other series?

Yes o   No o

c. In summer 2011, were you standardised online for any of your papers?

None o   Some o   All o

d. When did you first begin examining for AQA (or its predecessor Boards)? _____

e. In approximately how many examination series have you examined for AQA (or its predecessor Boards)?

_____ series.

f. In approximately how many examination series have you marked electronically?

_____ series.

g. For which other Awarding Body(s) have you examined either currently or in the past?
(Please tick all that apply)

OCR	In the past o	Currently o	CCEA	In the past o	Currently o
Edexcel	In the past o	Currently o	WJEC	In the past o	Currently o

Other, please state… _____.

h. Are you a…   Chair of Examiners   o   Chief Examiner   o
(please tick all that apply)   Principal Examiner   o   Examiner Team Leader   o
                               Assistant Principal Examiner   o   Assistant Examiner   o

i. Which qualifications, if any, did you examine during summer 2011? (please tick all that apply)

GCE (inc. Applied)   o   Key Skills   o   VRQ   o
GCSE (inc. Applied)   o   ELC   o   Other, please state… _____.
Diploma   o   Functional Skills   o

j. Which AQA office(s) did you deal with during the   Guildford   o   Harrogate   o
2011 examinations? (please tick all that apply)   Manchester   o

k. Are you …   …a full-time teacher   o   …a part-time teacher   o
               …retired   o   …other, please state…… _____.

l. If you teach, what kind of centre do you   Comprehensive School   o   F.E. College   o
work at?   Independent / Selective School   o   Other   o
           Sixth Form College   o   N.A.   o

m. Which components do you examine? _____.

n. Of these, which components are marked electronically? _____.

AQA