Handwritten *versus* typed essay exams: a question of equivalence

Liz Masterman with Jill Fresen
Technology Enhanced Learning
University of Oxford

www.digitaleducation.ox.ac.uk

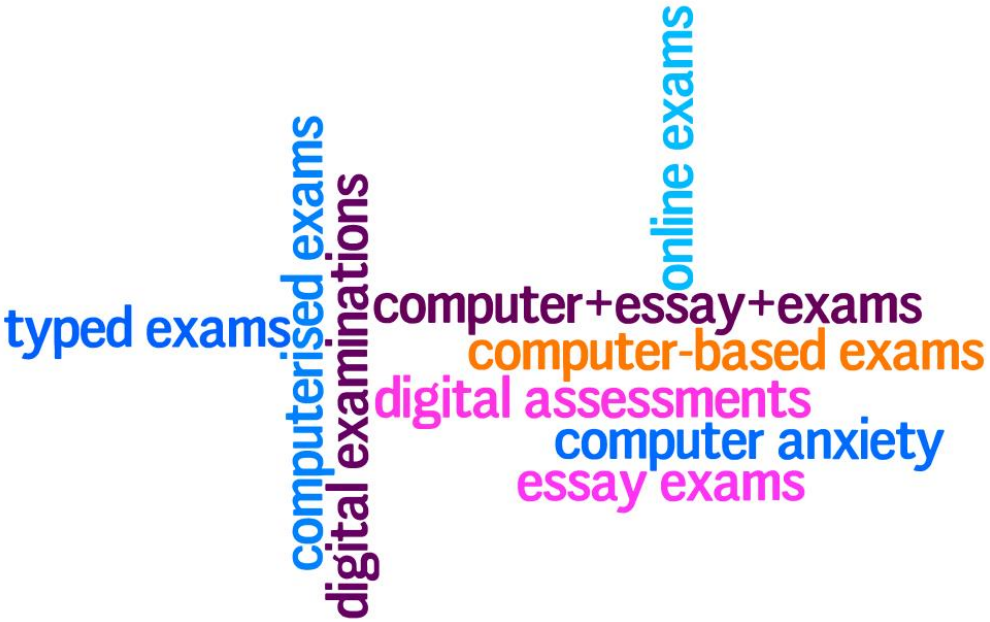
Are handwritten and typed exams equivalent?

'…the need for equivalence to be fully determined to ensure that overall performance outcomes are matched' (Noyes & Garland, 2008, p. 1357)

1. Students: differences in
   - Cognitive process?
   - Length and style of scripts?
2. Markers: differences in
   - Perception and attitudes re scripts?
   - Cognitive process?
3. Marks:
   - Is there a difference?
   - If yes, is it important?

Searching the peer-reviewed literature

- Google Scholar
- 2000 onwards + some older
- Followed up references and ‘related articles’ on journal websites
- 35 articles



The process of writing: a conceptual framework

- Planning
- Retrieving knowledge
- Translating ideas into text
  - Generating text
  - Transcription
- Revising
- Fluency in generating ideas and writing them down quickly frees cognitive resources to produce ‘reader-based prose’
- But exam environment can place pressure on these resources

(Peverley, 2006)

Students: sitting an e-exam (i)

- The process of writing
  - Self-reports by students
  - Inconsistent or contradictory data between studies
- The product: completed scripts
  - Length: typed = generally longer (various incl. Mogey et al., 2010)
  - Length of sentences; number and length of paragraphs: contradictory (various incl. Mogey & Hartley, 2012)
  - Lexical variation: typed = greater (Charman, 2014)
  - Lexical density: typed = less (Charman, 2014; Mogey & Hartley, 2013)
  - Style: typed = more informal (ibid., but disagreement e.g. Whithaus et al., 2008)

Students: sitting an e-exam (ii)

- Impact of change in tool
  - '… lack of fluency in lower order cognitive processes such as keyboarding or handwriting constrains higher order cognitive processes such as planning and reviewing.' (Kohler, 2015, pp. 140–141)

    i.e. typing proficiency has greater impact on performance than computer experience (e.g. Kohler, 2015; Mogey & Fluck, 2015)
  - But more proficient (faster) typists don't necessarily type more (Mogey & Hartley, 2010)
  - Anxiety re technical failure → increased pressure → more constraints on cognitive processes? (surmised by e.g. Hillier, 2014, but no data reported)
  - (Unacknowledged) inequities exist in handwritten exams too (Graham et al., 1998; Connelly et al., 2005)

Staff: perception of scripts

- Legibility:
  - Negative impact vs 'empathy' re poor handwriting (Lee, 2004; Powers et al., 1994)
  - Errors more visible in typed scripts (various incl. Kohler, 2015)
- Typed answers visually shorter (Powers et al., 1994)
  - Positive correlation between length of answer and mark awarded (various incl. Augustine-Adams et al., 2001)
- Possible expectation of higher quality in typed answers
  - 'Having the exams keyboarded seems to have shifted readers' expectations away from first-draft writing toward higher expectations associated with texts that have been more thoroughly revised' (Whithaus et al., 2008, p. 14; also various incl. Mogey et al.; 2012)

Staff: OSM and the process of marking

- Support for cognitive strategies (e.g. Shaw, 2008)
  - Skim-reading to extract salient themes and establish overall sense
  - Navigation within and between scripts
  - Both 'public' and 'private' annotation
- Impact on markers' performance
  - Severity: no significant difference (Whithaus et al., 2008; Johnson & various colleagues, 2009; 2012)
  - Accuracy: no significant difference (Johnson & various colleagues, 2009; 2012)
  - Reliability between markers: improved (Tisi et al., 2013)

Marks: is there a difference, and does it matter?

- Typed = higher: Augustine-Adams et al. (2001), Charman (2014), McCann et al. (2002) and Whithaus et al. (2008)
- Handwritten = higher: Bridgeman and Cooper (1998), Kohler (2015), Lee (2004), Mogey et al. (2010), Powers et al. (1994)
- Holistic scoring → handwritten higher
  Analytic scoring → typed higher
  (Lee, 2004; McCann et al., 2002)
- Differences not statistically significant…
- …but may matter on boundaries between grades
- Influence of research settings and methods

Implications: two views

‘The current findings do not indicate whether handwriting and keyboarding are significantly different cognitive processes or not.’
(Whithaus et al., 2008, p.17)

‘… the constructs measured in computer and paper modes are not the same. That is, the incorporation of computers into writing assessments involves a new way of thinking about composing processes … Inevitable sources of non-equivalence of the construct … might lead to differences in test performance to some extent.’
(Lee, 2002, p. 152)

Implications

'… the constructs measured in computer and paper modes are not the same.' (Lee, 2002, p. 152)

- Rule out choice between handwriting and typing (or is students' right to choose more important?)
- Provide opportunities for students to develop typing proficiency (or should this be their responsibility rather than the institution's?)
- Rule out choice between marking on paper and on screen (or is academics' freedom to choose more important?)
- Further research: e.g.
  - Use keystroke tracking to triangulate students' self-reports
  - More comparative analysis of scripts
  - Investigate markers' cognitive strategies OSM vs paper

A trial of e-exams at Oxford

- 20-21 April
- Inspera software
- 4 exams
  - Formative: Theology, Computer Science
  - ‘Mocks’: 2 modules in MSc Law & Finance
  - >70 undergraduate and taught master’s students
- Typed only: no option to handwrite
- Equipment: BYOD and/or University-supplied (per-exam basis)
- Evaluation in progress