

Editorial

Victoria Crisp

Welcome to the autumn issue of *Research Matters*. Disruption to school life due to emergency situations can have a substantial effect on students in terms of their learning and wellbeing. There is a need for schools to develop strategies for addressing the challenges of such situations. Our first article explores this issue in the context of the COVID-19 pandemic. Through a detailed analysis of interviews with teachers, Filio Constantinou identifies a set of macro- and micro-strategies used by schools and discusses the broader implications for emergency readiness.

In our second article, Tom Benton addresses a decision that has to be made about any high stakes test for any qualification type and subject: how long the test should be. Tom looks at the recommendations for minimum levels of reliability from the literature and uses these and psychometric formulae for the relationship between reliability and test length to calculate possible recommendations for test durations. To provide perspective, these recommendations are then compared to the test durations used in practice in a number of assessment contexts in different jurisdictions.

The importance of maths to everyday life and career opportunities continues to motivate moves to encourage greater uptake of maths among 16- to 18-year-olds. In our third article, Tim Gill looks at one example of this: Core Maths qualifications. These qualifications, which were first assessed in 2016, are targeted at learners who achieve at least a grade 4 in GCSE Maths but who do not go on to study A Level Maths. Tim's analysis explores the background characteristics of those taking Core Maths, the other qualifications and subjects these learners also study, and whether taking Core Maths is associated with better results in other qualifications with a quantitative element.

Moves towards increased use of digital examinations in general qualifications raise many interesting issues in relation to comparability. One important theme is how typing or word-processing responses, instead of handwriting them, may affect comparability. In our fourth article, Santi Lestari reviews the existing literature on the comparability of typed and handwritten long-answer responses, in terms of scores, text characteristics, marking, and composing processes.

In our final article, Lucy Chambers, Emma Walland and Jo Ireland describe a study in which they explored the use of the Comparative Judgement method for the assessment of music compositions and performances based on audio recordings. This is a novel use of Comparative Judgement as the method has primarily been applied to learners' written work. Lucy and colleagues use questionnaire data to investigate how judges felt about comparing music audio recordings and the factors influencing decision-making in this context.