Whole-class questions in the active classroom: digging deeper

Ross K. Galloway, Tom Summers

ross.galloway@ed.ac.uk @RossKGalloway

EdPER Edinburgh Physics Education Research

THE UNIVERSITY of EDINBURGH

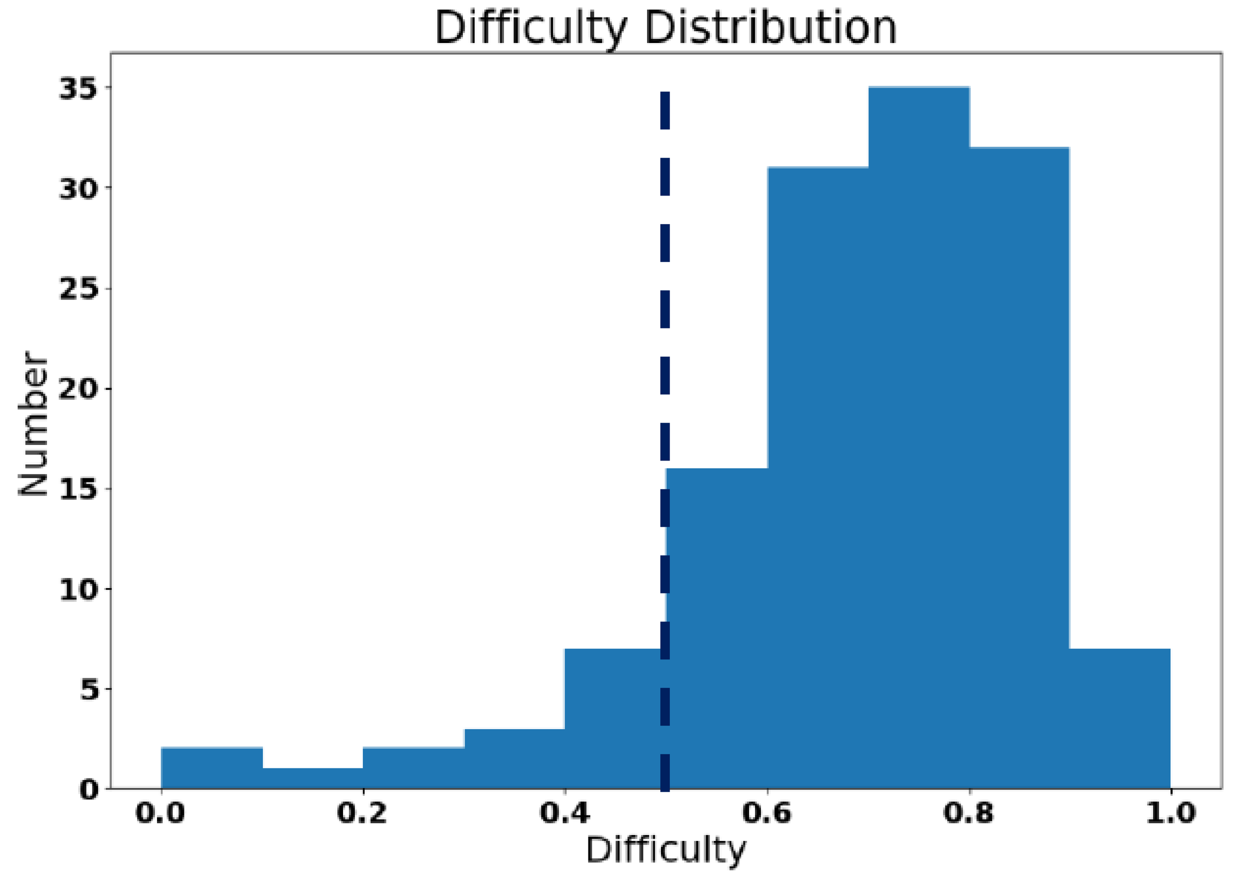
1. Motivation

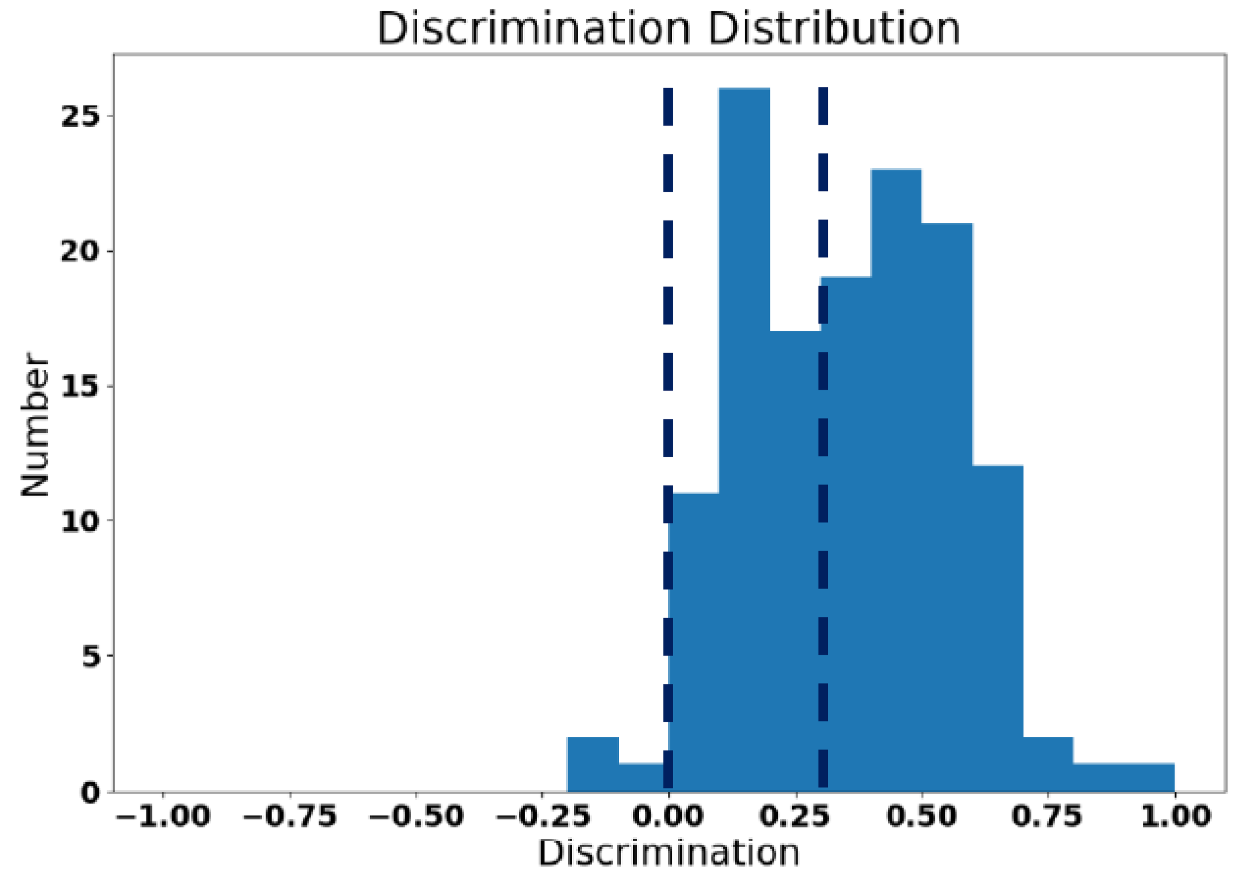
In-class electronic voting ('clicker questions') have been used for over a decade in our flipped-class introductory modules. The effectiveness of these questions can be judged on an item-by-item basis, but we wished to investigate more systematic means of evaluating their holistic performance.

Classical Test Theory [1] provides a standard, well-understood suite of statistical evaluations, but is more commonly applied to diagnostic instruments that are designed to test relatively few underlying concepts. Can it be usefully applied to the much more heterogeneous population of clicker questions used over a whole semester? We investigated this issue.

In addition, Principal Component Analysis [2] is a standard technique for identifying any underlying correlations in a set of questions; we subjected our clicker questions to a standard PCA investigation.

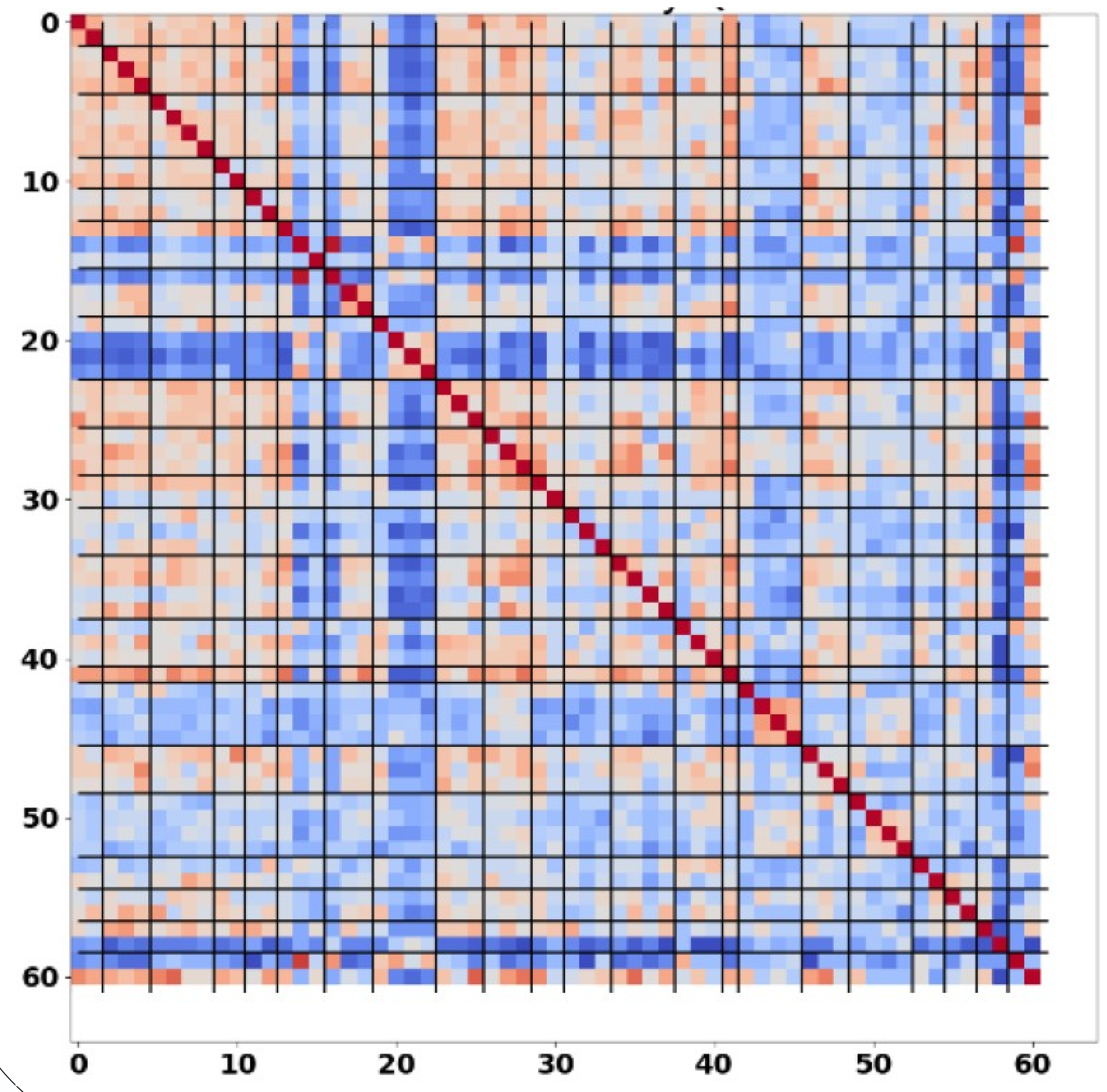
2. Classical Test Theory – selected findings





Distribution of clicker question difficulty indices (left): counterintuitively, under the standard definition, a higher 'difficulty index' corresponds to an easier question. The optimum difficulty index is 0.5, so we see that our clicker questions are generally undesirably easy. Distribution of clicker question discriminatory power (right): a value of 0.3 or greater is desirable, so a number of questions discriminate relatively poorly between students with stronger and weaker grasps of the concepts. This is related to the previous finding: questions that are too easy discriminate poorly since most students get them correct. However, pleasingly, almost no questions have a negative discrimination index: these are questions that weaker students answer correctly more often than generally stronger students (indicating a pathological or malfunctioning question.)

3. Principal Component Analysis – selected findings



Correlation matrix of clicker question correctness (left) over a suite of 60 questions delivered over one semester. More intense colouration indicates greater correlation (red for positive, blue for negative). The black grid indicates boundaries between discrete lectures. As would be anticipated, questions tend to be most strongly correlated with other questions in the same lecture, since individual lectures typically address congruent topics. However, many inter-lecture correlations are also seen. PCA was used to identify correlated sub-sets of questions.

Some components were obvious, e.g. questions all relating to the Centre of Mass. Some were more subtle but not unexpected, e.g. superficially unrelated questions that all required decomposition of vectors into orthogonal components. However, some components were not obvious at all, e.g. a set of 6 completely unrelated questions that all loaded on to the same component as did a diagnostic question from the Cognitive Reflection Test [3]. The CRT measures ability to suppress instinctive responses and activate deeper cognition; on closer inspection, all these disparate questions had 'obvious' but incorrect solutions that must be disregarded. This suggests that the CRT can be a useful wider indicator for effective thinking in physics.

4. Implications for practice

Our findings from Classical Test Theory support its effective use in understanding the behaviour of in-class clicker questions: its standard statistical tools allow the identification of well-functioning and pathological questions. Instructors can use the standard tests (available in most statistical packages) 'off the shelf' to usefully analyse their clicker questions.

Our use of Principal Component Analysis has identified meaningful correlated components in the question responses, some of which would not be obvious from manual inspection of the data. This can yield unexpected insight into the answer profiles. For example, it has provided evidence for a connection between inhibitory control within 'trivial' problems (in the CRT) and inhibition in more complex situations of physical reasoning. This is consistent with findings in cognitive neuroscience about the fundamentals of scientific reasoning [4].

References

[1] P.V. Engelhardt, *An Introduction to Classical Test Theory as Applied to Conceptual Multiple-choice Tests*, PER-Central (2009)

[2] L. Ngo *Principal component analysis explained simply*, BioTuring (2018)

[3] S. Frederick, *Cognitive reflection and decision making*, J. Econ. Perspect. **19**, 4 (2005)

[4] L. Nenciovici, G. Allaire-Duquette & S. Masson, *Brain activations associated with scientific reasoning: a literature review*, Cognitive Processing **20**, 139 (2019)