Using data to improve examinations and learning in Hong Kong: the HKEAA experience

Dr. Neil DRAVE
Senior Manager (Assessment Development - English)
Hong Kong Examinations and Assessment Authority (HKEAA)

HKEAA develops and administers the Hong Kong Diploma of Secondary Education (HKDSE) examination, which serves to certify school leavers at the end of six years of secondary school education and is also Hong Kong's main qualification for post-secondary admission.[1] In response to stakeholder requests for greater transparency about its operations, and to better support teaching and learning, HKEAA has been expanding the range of the post-examination information it provides to candidates, schools and teachers. Examination data are also used within the HKEAA for grading purposes and to review the reliability of examination papers, and by the Hong Kong Government to investigate the influence of socioeconomic factors on student performance, but these uses are beyond the scope of this paper.

The following kinds of post-examination information are provided to stakeholders:

- Hard copy *Question Papers* booklets, with marking schemes and comments on candidates' performance, available to the general public.
- Samples of performance at different levels of attainment (Level 1 to Level 5), some of which are annotated, posted on the Authority's web site.[2]
- Post-examination briefing sessions for teachers.
- A *Moderation Report* (for those subjects with a School-based Assessment component).
- *Examination Reports* on the whole cohort's performance in general and in each subject, available to the general public.
- Reports on the examination performance of particular schools, produced on request and for a fee (*School Statistical Report*).

The Authority also provides access to an online examination analysis platform (the *Assessment Quality Platform*, *AQP*) which can be used by schools for a fee to analyse their own test data, and offers assessment literacy courses to help teachers understand HKEAA operations and interpret data.

It is difficult to know the impact of providing most of these information types since we do not know who accesses them, but it *is* possible to track users of School Statistical Reports since there is a record of purchasers. We therefore survey schools periodically to evaluate the provision of these Reports. The rest of this paper reports on the results of the 2018 survey, concluding with general reflections on data and learning.

School Statistical Reports

The HKDSE School Statistical Report is a tailor-made report for individual schools. It allows users to understand their students' performance in the HKDSE and compare it with the whole school candidature in a given year. Item marks captured from an onscreen marking system are the basis for the statistical analysis in the Report.

There are three kinds of Reports:

- School Report: this gives data on all candidates from a single school.
- Class Report: this gives data on a single class of candidates.

[1] See Cheung (2016) for a comprehensive overview of HKEAA's work; also http://www.hkeaa.edu.hk/en/hkdse/introduction/

[2] Most HKDSE subjects employ standards-referenced reporting of results.

- Teaching Group Report: a teaching group is a unit that combines students from different classes.

Reports are available in hard copy and as PDF files on CD, and are published around four months after the release of examination results. They can be purchased separately but all schools in the 2018 survey had opted to receive at least the School Report.

School Reports have the following sections:
1.1 Results Summary
1.2 Statistics on general performance
1.3 General performance in the best five subjects
1.4 Statistics for eligibility to sub-degree programmes / civil service appointments
1.5 Statistics related to university admission
1.6 Grade point distribution in best 5 subjects (statistics related to university admission)
2 Category A subject results[3]
3 Item analysis for onscreen marked subjects
4 Multiple-choice item analysis

The Reports exemplify the kinds of 'large data' (as opposed to 'Big Data') which are routinely made available by examination bodies following summative examinations.

User survey 2018
A 15-item survey containing both closed- and open-response questions was administered via the HKEAA e-survey web site. The purpose of the survey was to find out whether schools were satisfied with the quantity and types of data provided in the Report, and to get a general sense of how the data were used.

- Number of schools targeted: 366 schools which received 2017 School Reports[4]
- Number of individual responses: 378 (because respondents in different roles in the same school were allowed to respond separately from their own perspective)

Almost all responses were in English (the second language of nearly all respondents) and the few which were in Chinese were ignored for this research. The responses are summarised below.

Findings
About 64% of respondents were Panel Chairpersons (PC)[5], 21% were teachers, 14% were Vice Principals (VP) and the remainder Principals (P). The highest proportion of respondents were from the Mathematics subject area (14.3%), followed by Chinese (8.5%) and Liberal Studies (7.4%), which is understandable given that these are compulsory subjects.

1. Report format (Question 1)
Close to two thirds of respondents reported receiving the Report on CD only, close to one third received both formats, and the remainder (8.7%) received hard copies only.

2. Access (Questions 2 and 3)
- Hard copies were kept in the office of the VP or P, according to about 40% of respondents; about a quarter in total said that the Report was kept it in the school office, library or staff room.

[3] Category A: elective subjects. Candidates usually study two or three of these, in addition to the four compulsory subjects of Chinese, Maths, English Language and Liberal Studies.
[4] There are about 500 secondary schools in Hong Kong.
[5] PCs are department heads i.e. lead a team of teachers in a particular subject. The Principal is in overall charge of a school but usually does not teach.

- Access was open to P, VP, the subject Panel Chairperson (PC) and teachers but in some cases certain roles had access to (or used only) certain kinds of data, e.g. the PC worked only with the parts of the Report on their own subject.

3. Users (Question 4)

- The PC used the reports most often, followed by teachers, VP and P. Very few other users were named. One caveat to note is that 'use' was not defined in the survey and could therefore have been interpreted in different ways.
- There was an even split between using the report when it arrived (November) and when planning the school year; 9% used it before the exam period (which takes place in April and May).

4. 'How do the different people use the Report?' (Question 5)

In this open-response question, respondents expanded upon their answer to Question 4.

- The vast majority of respondents, no matter what their position, reported making use of the data.
- Responses were quite general overall, with infrequent detailed explanations, making it difficult to identify particular patterns of use.
- PC used the Reports most; they reporting passing on information and giving advice to teachers in their team. Note that most respondents were PCs, however.
- Differentiation according to role:
    - P/VP: these roles were most often mentioned in the context of getting an 'overall picture/overview' of the data in order to allocate resources or pass on information to others.
    - PC: engaged in 'cognitive processes' (Halliday 1985) such as 'digest[ing] the data' and 'evaluat[ing] subject performance' in order 'to help' with the 'behavioural processes' (Halliday 1985) of 'adjust[ing] teaching', 'modify[ing] strategies' and sharing information with teachers.
    - PC/teachers: used data to (help) revise teaching plans/strategies, devise test items in internal assessments, and for drilling students.
- An analysis of word frequencies gives the following insights:
    - 'Teaching' was mentioned more than twice as often as 'Learning' (73 vs. 30 occurrences) and almost all occurrences of 'learning' co-occurred with 'teaching'.
    - 'Weaknesses' (38 occurrences of the lexeme 'weak') were more prominent than 'strengths' (22 lexeme occurrences).

5. Report usefulness (Questions 6, 7 and 8)

- The overall average usefulness rating was 4.88 out of 6, suggesting that the Reports were felt to be useful.
- The most useful sections overall were Item Analysis, followed by MC Item Analysis and comparison with other schools. Least useful were statistics related to eligibility for sub-degree programmes and the civil service.
- The most useful report section for Principals was 'general performance in best 5 subjects'.
- The highest ratings on the measure 'usefulness for improving teaching & learning' were given by subjects concerned with numbers, such as Maths and Economics, while arts and humanities subjects were less positive.
- The amount of information was felt to be appropriate ('the right amount of information' 93%; 'too much information' 5%), as was the type(s) of information provided ('appropriate' or 'mostly appropriate' 98%).

6. 'In what ways is the Report helpful or not helpful for improving teaching and learning?' (Question 9)

The purpose of this question was to obtain more detail about how the Reports are used to inform teaching and learning.

- As in Question 5 above, responses were lacking in detail and tended to rely on stock phrases to describe what was done (e.g. 'analyse performance').
- Most often, greater understanding was the intended outcome.
- Reports allowed teachers to perform mental processes (e.g. 'know', 'evaluate', 'understand') which in turn led to the behavioural processes of 'adjust[ing]' 'focus[ing] on', 'modify[ing]', 'designing' and 'reviewing' teaching 'tools', 'activities' and 'strategies'. Reports helped teachers to help students rather than helping students directly.
- There were many occurrences of the modal verb 'can' (e.g. 'can help', 'can know', 'can modify'). It is difficult to explain this, however, since 'can' has many functions.
- Students were mentioned often but almost always only in terms of their 'performance', 'weaknesses' or 'errors'.
- Strengths and weaknesses were often mentioned together but when only one of these was mentioned, it was weaknesses.
- The practical steps for changing teaching practice mentioned most often were the allocation of more time and more drilling.

Conclusions

It seems that respondents were generally satisfied with the Reports and felt that they contained the right amount of the right kinds of information. They valued data on specific items the most and used them to think about changing their teaching to address students' areas of weakness. Exam data exposed student weaknesses, which led to a strategy to address them, for example, by allocating more time and through drilling.

Respondent comments suggested a positive, constructive appreciation of exam data as part of a considered and targeted strategy for improvement. This was particularly true of responses from subjects related to numeracy. This is understandable given the expertise of the teachers, as well as the ways these subjects are tested: Maths, for example, is largely item-based while other subjects (such as English Language) use writing tasks, meaning that statistics give less information about the quality of a candidate performance. There were some negative and cynical comments, but these were few and far between.[6]

Although this survey did not reveal much about the specifics of classroom practice, perhaps because most respondents were PCs, it suggests that exam data do have a role to play in the strategic planning of teaching. This finding is in accord with work by Brown et al. (2009) which suggests that Hong Kong teachers value examinations and conceive of exam preparation as a means of improving learning outcomes (by making students accountable).

Discussion

Examination bodies have a responsibility to provide reliable, useful and meaningful data to stakeholders. It seems that the 'large data' format discussed above is well-received by schools and is being used as intended. These are undoubtedly positive factors. There are outstanding questions for HKEAA, however, such as the following:

- Whether exam results are *actually* improved by reviewing examinations data. While it is possible to tell whether schools perform better or worse from year to year, it is difficult to isolate the influence of exam data. The emphasis on comparison with other schools might also run counter to the philosophy underpinning the HKDSE, standards-referencing.[7]
- Which subjects/papers benefit most from exam data. The survey responses suggest that data are most welcomed by Maths-related subjects: do these subjects benefit most?

[6] Respondents had paid for the Reports (although not personally), and knew that they were doing an HKEAA survey, which might have influenced the results.

[7] Hong Kong initiated the move from norm-referenced public examinations to standards-referencing in 2007, with full implementation of standards-referenced reporting in the HKDSE, first held in 2012.

- Whether examination data correlate positively with 'Factors affecting student achievement' (Hattie's 2008); 'planning and predicting' and 'estimating student achievement' are two factors which might be affected, although the latter most likely refers to formative rather than summative processes.
- What specific teaching strategies, if any, are informed by exam data: this will be explored in a follow-up study.
- Whether HKEAA should train stakeholders to interpret Reports. Training is currently provided only to AQP users, not Report purchasers, and it is possible that some users incorrectly interpret the data, or misuse it.
- Whether HKEAA should provide non-statistical data. Harvard School of Education's 'Data Wise' project and *Driven by Data* (Bambrick-Santoyo 2010) exemplify how data can be used systematically for academic improvement. These approaches analyse exam data, but also other information sources such as samples of student writing and homework assignments. HKEAA examination scripts take the form of scanned handwritten answers, so it is not currently feasible to use these for large-scale research, and they are not provided to schools. Whether to provide them to schools, and if so in what format, needs careful consideration.

There are other matters which are beyond the scope of HKEAA's work but which schools and planners need to consider:

- Whether teacher competence and morale are influenced by the provision of exam data. While teaching group level examination reports such as those provided by HKEAA (but not the focus of this paper) can give an insight into the exam outcomes of students taught by particular teachers, they cannot tell us anything else about the teachers.
- Whether students' needs, motivation and self-direction are influenced by teacher use of exam data: in this study, students were seen in terms of weaknesses to be improved, but approaches such as Assessment for Learning (AfL) and Assessment *as* Learning, both elements of Hong Kong education policy, include other, affective factors (Li and Wu 2018).
- *Why* students can or cannot do certain things. Exam data are by nature descriptive rather than analytical, although they can be the basis for analysis, and their summative nature means that any analysis can help only the *next* cohort of students. In the data driven approach cited above, interim (e.g. quarterly) assessments are preferred over annual assessments in being timelier: they allow one to (metaphorically) cure the patient while they are alive rather than wait for the autopsy (Bambrick-Santoyo 2010: 137). They are also considered to be more useful than 'in-the-moment' assessments (such as AfL) because they give a better sense of an end goal to be worked towards.
- Whether matters such as school-wide administration, and 'soft' educational factors like happiness and school ethos, are affected. While exam data are useful for comparing school results to Hong Kong as a whole, they cannot tell schools what changes might improve them as organisations.

There are also broader issues relating to the possible unintended consequences of providing exam data:

- Providing data may have the long-term consequence of narrowing the focus of teaching to exam preparation *even more* than the existence of exams alone does, especially in places (such as Hong Kong) where examinations are perceived to be a positive force for motivating students, and where 'success' often equates to 'exam success'.
- Schools have to be careful that that they do not focus so much on exam data that they neglect other, (more) productive, long-term educational goals. Leung's (2014) case study of leadership and educational change in two schools is an example of what can go wrong in this regard.
- Providing *summative* exam data may make *formative* assessment difficult to implement. Brown et al. (2009) suggest that there will be resistance to formative assessment *precisely because* summative examinations are seen as a positive force and get so much attention.

- When an organisation perceived to be an authority on learning outcomes (e.g. an examinations body) provides data of a certain kind in a certain format, this may give rise to the false impression that such data are infallible. We need to remind ourselves and stakeholders that data can never be completely comprehensive, trustworthy, objective or context-free; and any patterns which can be detected are artifacts of analysis rather than facts (Boyd and Crawford 2012).

One can argue about the relative responsibilities of examination bodies, governments and schools for the effective dissemination and use of data. What is clear, however, is that if bodies such as HKEAA are to make a positive contribution to educational outcomes, we need to take a reasoned and ethical approach to data provision, making sure that data serve learning and not the other way round. Stakeholders need to be encouraged to use the full range of information available to them and to be cautious when interpreting data.

Big Data assessment, such as computer-mediated testing and algorithm-driven approaches, may be helpful in addressing, or at least bypassing, some of these 'large data' issues. It is likely to more flexibly assess a wider range of skills and behaviours, changing the way we define 'examinations' and even 'data'. As Boyd and Crawford (2012) point out, however, such assessments will pose new challenges; for example, they will be constrained by the available technology, and their broad reach is likely to give rise to serious ethical considerations. Meaningful discussion about how best to use 'large data' should help us prepare for the advent of Big Data.

The author would like to thank William Lai and colleagues from the HKEAA Examination Systems Unit for their help with administering the survey and summarising the quantitative findings. The opinions expressed above are the author's own and do not represent the official position of the HKEAA.

References

Bambrick-Santoyo, Paul. 2010. *Driven by Data: A Practical Guide to Improve Instruction*. San Francisco: Jossey-Bass.

Boyd, Danah and Crawford, Kate. 2012. 'Critical questions for Big Data.' *Information, Communication & Society*, 15:5, 662-679. DOI: 10.1080/1369118X.2012.678878

Brown, Gavin, Fok, Ping Kwan, Kennedy, Kerry and Yu, Wai Ming. 2009. 'Assessment for improvement: Understanding Hong Kong teachers' conceptions and practices of assessment.' *Assessment in Education*. https://www.researchgate.net/publication/233871183

Cheung, Thomas. 2016. *The Management of Public Examinations in Hong Kong*. Hong Kong: HKEAA.

Di Carlo, Matthew. July 2010. 'Teachers matter, but so do words.' http://www.shankerinstitute.org/blog/teachers-matter-so-do-words

Halliday, MAK. 1985. *An Introduction to Functional Grammar*. London: Edward Arnold.

Hattie, John. 2008. *Visible Learning*. London: Routledge. Updated list of factors: http://www.evidencebasedteaching.org.au/hatties-2017-updated-list/

Lai, Wai Lam William. August 2016. 'How assessment data can enhance teaching and learning.' Paper presented at the IAEA Conference, Cape Town, South Africa.

Leung, Lai Sum. 2014. 'A case study of educational change and leadership in Hong Kong primary schools.' *Hong Kong Teachers' Centre Journal*, 13, 97-123. https://www.edb.org.hk/HKTC/download/journal/j13/C01.pdf