

IAEA Conference 2018

Prediction of School Candidates' Performance in the Hong Kong Diploma of Secondary Education Examination

LEE WONG Wai Christina, WEN Zhong-lin, CHANG Rui

Hong Kong Examinations and Assessment Authority

Abstract

The Hong Kong Diploma of Secondary Education Examination (HKDSE) was first administered in 2012 as part of the education reform in Hong Kong. It aims to measure the attainment of students upon their completion of the 3-year senior secondary curriculum. Standards-referenced Reporting (SRR) is adopted to ensure that examination results are transparent and explicit. Candidates' achievement is reported against a set of fixed standards divided into five levels, with 5 being the highest. Descriptors are written for each subject to describe what a candidate performing at a given level is typically able to do. Samples of candidates' performance are also provided to illustrate the standards.

In order to study whether schools understand the standards and can make use of the level descriptors to predict their students' public examination results accurately, the HKEAA has been conducting a study on the reliability of schools' predictions of their students' HKDSE attainment levels in the four core subjects - Chinese Language, English Language, Mathematics and Liberal Studies. A good understanding of the standards will facilitate teaching and learning, and accurate predicted levels will help teachers manage the expectations of students and parents and enable them to make realistic decisions on possible pathways for work and study.

Schools are required to submit the scores and the predicted levels of each candidate for the four core subjects based on internal examinations. The data are compared against the actual scores and levels achieved by these candidates in the HKDSE. Feedback is provided to schools at the start of the following academic year with regard to their over or under-estimation.

A pilot study was conducted in 2013 with 24 schools. The number of participating schools increased to 165 in 2017, and data were collected from about 20,000 candidates. A longitudinal analysis was also carried out in 2017 for the 59 schools that participated in the study for four consecutive years to see whether improved familiarity with the HKDSE standards had resulted in more accurate predictions across years.

Background

A new academic structure was implemented in Hong Kong in 2009 as part of the Educational Reform which started in the year 2000. The proposal for the reform and its aims and scope were laid down by the Education Commission in 2000 (Education Commission, 2000). The major changes include curriculum reform to specify core knowledge and higher order skills for all students, criterion-based assessment aligned to the curriculum and an exit examination for both certification and matriculation purposes.

After completing a three-year junior secondary curriculum, all students can continue to study a three-year senior secondary curriculum. The new senior secondary academic structure provides all students with the opportunity to receive six years of secondary education to promote life-long learning and to meet the changing societal needs of Hong Kong. The Hong Kong Diploma of Secondary Education Examination (HKDSE) was first administered in 2012 to replace the previous public examinations, namely the Hong Kong Certificate of Education examination (HKCEE) and the Hong Kong Advanced Level examination (HKALE), which were taken at the end of five and seven years of secondary schooling respectively. The new secondary curriculum has been shortened to six years with a corresponding change in the length of degree programmes from three to four years in the tertiary sector. The HKDSE serves the dual purpose of providing certification of students' attainment in their chosen subjects, and also as an instrument for selection to further studies.

Standard-referenced Reporting

Standards-referenced reporting (SRR) is adopted for the HKDSE. SRR aims at reporting candidates' results against a set of prescribed standards based on typical performances of candidates at those levels. The results are expressed in terms of five levels of performance, of which 5 is the highest and 1 the lowest. In order to facilitate finer discrimination at the top end, norm-referencing is applied for the Level 5 candidates: with the best performance awarded a 5**, and the next top group awarded a 5*. A performance below Level 1 will be labelled as 'Unclassified'. In setting the standards, rigorous procedures are adopted to ensure that the standards are appropriately established and maintained (HKDSE Grading Procedures and Standards-referenced Reporting Booklet 2018).

For each of the levels, descriptors have been developed describing what the typical candidate performing at this level is able to do. The descriptors were developed by

subject experts including experienced markers, teachers, principals and curriculum experts with reference to HKCEE and HKALE standards. To ensure that there is no mismatch between students' abilities and the anticipated standards, sample papers were piloted on senior secondary students with different abilities from a range of schools. These sample papers also served the purpose of illustrating the specific requirements of the examination to candidates. The performance samples collected were analysed and used as the basis for the development of level descriptors which illustrate the standards expected at the various levels of performance. Level descriptors are important reference sources for subject experts to make judgments on grading and also facilitate tertiary institutions and employers to set appropriate admission and job requirements.

In an effort to better promulgate the HKDSE standards, SRR Information Packages for individual subjects were published and distributed to schools in mid-2009. Each package contains the assessment framework of the subject, the level descriptors, sample papers and suggested answers / marking guidelines, exemplars of typical candidate performance. After each examination, backdated question papers are also published with the marking schemes and examiners' comments on candidates' performance. Samples of candidates' performance and other relevant information are available on the HKEAA website (HKDSE Subject Information 2018).

The Level Prediction Study

One of the advantages of SRR is that end-users of the HKDSE qualification are provided with more explicit information about the attainment of candidates than for the HKCEE and HKALE. Because the level descriptors spell out the standards required for different levels of performance which are also illustrated by exemplars of actual candidates' work, more information is available to students and teachers on the requirements of the HKDSE which may help to improve learning and teaching. However, despite the information provided, it is not certain whether schools have a good understanding of the HKDSE standards, especially in the initial years of implementation. It was therefore decided that a study be conducted to estimate the reliability of school predictions on the subject results of their students.

Only the four core subjects, namely Chinese Language, English Language, Mathematics (Compulsory Part) and Liberal Studies (CEML) are included in the study as they are taken by all students and the sample size is adequate. Another important factor is that universities in Hong Kong have set a threshold of levels 3322 in CEML as the minimum requirement for admission to degree programmes. Reliable

predicted subject results will help schools to manage the expectation of students and parents and are useful for helping them make decisions on possible future pathways for work and study before the release of public examination results.

Data collection and Preparation

Schools are invited to take part in the research study and provide their final examination subject marks and predicted levels of their students in the core subjects. Feedback is given to the schools after the release of the rechecking and remarking results. The data collected are compared against the actual scores and the actual levels obtained by these candidates in the HKDSE. The full marks (maximum possible marks) of the school examination for the four subjects are also collected. For each subject, the scores obtained from school are standardised by using the full mark to ensure comparability across schools. Students with zero scores in school examinations or who are absent for some examination papers are excluded when analysing the relationship between school results and HKDSE results.

Analyses Conducted and Feedback to Schools

The following analyses are conducted:

1. Level comparison
 - Cross-tabulation
 - Discrepancy distribution
 - Kappa statistics
 - Regression analysis
2. Score comparison
 - Global regression
 - Sub-group regression

A report is prepared for each participating schools which provides the following information:

1. Comparison between the predicted level and actual level for CEML
2. Discrepancy between predicted level by school and the actual HKDSE level
3. Correlation between the scores from school and actual scores for CEML

A briefing session is held after the study each year at which the reports are distributed and explained to participants. Representatives from schools with good predictions are invited to share their experience with other schools on various aspects of assessment

and teaching. A survey is conducted at the end of the session to collect comments and suggestions from participants.

Results of the 2017 Study

The data of 19,867 candidates from 165 schools were received and analysed. For each of the Core Subjects CEML, the score and the predicted level for each candidate were obtained from schools before the release of HKDSE results.

Levels comparison

Cross tabulation – consistency between predicted and actual levels

The predicted levels and the actual HKDSE levels of the candidates in each of the core subjects were compared. The cross tabulation of predicted level and actual HKDSE levels for the four subjects are shown in Table 1 to Table 4.

Table 1. Cross tabulation of Actual HKDSE Level and Predicted Level for Chinese

| | | Actual level (Chinese) | | | | | | | | Subtotal |
|--|------------|------------------------|------|------|------|------|------|-----|-----|----------|
| | | U | 1 | 2 | 3 | 4 | 5 | 5* | 5** | |
| Predicted level U (Chinese) | U | 148 | 162 | 49 | 5 | 1 | 0 | 0 | 0 | 365 |
| | 1 | 147 | 1024 | 944 | 108 | 12 | 0 | 1 | 0 | 2236 |
| | 2 | 21 | 548 | 3085 | 1560 | 279 | 23 | 4 | 0 | 5520 |
| | 3 | 3 | 25 | 1290 | 2732 | 1373 | 165 | 38 | 5 | 5631 |
| | 4 | 1 | 4 | 190 | 1262 | 1759 | 543 | 158 | 23 | 3940 |
| | 5 | 0 | 0 | 11 | 155 | 525 | 359 | 206 | 49 | 1305 |
| | 5* | 0 | 0 | 3 | 26 | 115 | 140 | 154 | 69 | 507 |
| | 5** | 0 | 0 | 1 | 0 | 15 | 32 | 39 | 53 | 140 |
| Subtotal | | 320 | 1763 | 5573 | 5848 | 4079 | 1262 | 600 | 199 | 19644 |

Table 2. Cross tabulation of Actual HKDSE Level and Predicted Level for English

| | | Actual level (English) | | | | | | | Subtotal | |
|--|------------|------------------------|------|------|------|------|------|-----|----------|-------|
| | | U | 1 | 2 | 3 | 4 | 5 | 5* | | 5** |
| Predicted level U (English) | U | 822 | 142 | 6 | 1 | 1 | 0 | 0 | 0 | 972 |
| | 1 | 510 | 1265 | 421 | 12 | 1 | 2 | 0 | 0 | 2211 |
| | 2 | 25 | 705 | 3247 | 656 | 15 | 1 | 1 | 0 | 4650 |
| | 3 | 0 | 17 | 984 | 3681 | 825 | 17 | 2 | 0 | 5526 |
| | 4 | 0 | 0 | 23 | 925 | 2567 | 449 | 43 | 0 | 4007 |
| | 5 | 0 | 0 | 0 | 21 | 530 | 650 | 260 | 26 | 1487 |
| | 5* | 0 | 0 | 0 | 1 | 41 | 211 | 287 | 92 | 632 |
| | 5** | 0 | 0 | 0 | 0 | 0 | 17 | 81 | 120 | 218 |
| Subtotal | | 1357 | 2129 | 4681 | 5297 | 3980 | 1347 | 674 | 238 | 19703 |

Table 3. Cross tabulation of Actual HKDSE Level and Predicted Level for Mathematics

| | | Actual level (Mathematics) | | | | | | | Subtotal | |
|--|------------|----------------------------|------|------|------|------|------|-----|----------|-------|
| | | U | 1 | 2 | 3 | 4 | 5 | 5* | | 5** |
| Predicted level U (Mathematics) | U | 548 | 184 | 27 | 6 | 0 | 0 | 0 | 0 | 765 |
| | 1 | 413 | 1096 | 760 | 46 | 4 | 0 | 0 | 0 | 2319 |
| | 2 | 48 | 501 | 2709 | 1255 | 179 | 1 | 0 | 0 | 4693 |
| | 3 | 3 | 22 | 714 | 2252 | 1513 | 51 | 0 | 0 | 4555 |
| | 4 | 0 | 2 | 48 | 626 | 2989 | 713 | 85 | 4 | 4467 |
| | 5 | 0 | 0 | 4 | 8 | 618 | 796 | 306 | 37 | 1769 |
| | 5* | 0 | 0 | 0 | 1 | 67 | 333 | 341 | 114 | 856 |
| | 5** | 0 | 0 | 0 | 0 | 1 | 21 | 120 | 133 | 275 |
| Subtotal | | 1012 | 1805 | 4262 | 4194 | 5371 | 1915 | 852 | 288 | 19699 |

Table 4. Cross tabulation of Actual HKDSE Level and Predicted Level for Liberal Studies

| | Actual level (Liberal Studies) | | | | | | | | Subtotal |
|----------------------------|--------------------------------|------|------|------|------|------|-----|-----|----------|
| | U | 1 | 2 | 3 | 4 | 5 | 5* | 5** | |
| Predicted level U | 126 | 183 | 66 | 14 | 2 | 0 | 0 | 0 | 391 |
| (Liberal Studies) 1 | 105 | 719 | 755 | 252 | 51 | 2 | 0 | 0 | 1884 |
| 2 | 5 | 373 | 2016 | 1703 | 377 | 11 | 0 | 0 | 4485 |
| 3 | 0 | 44 | 993 | 2983 | 1812 | 106 | 27 | 1 | 5966 |
| 4 | 0 | 3 | 109 | 1236 | 2870 | 477 | 150 | 21 | 4866 |
| 5 | 0 | 0 | 6 | 94 | 747 | 332 | 160 | 43 | 1382 |
| 5* | 0 | 0 | 0 | 17 | 194 | 159 | 148 | 59 | 577 |
| 5** | 0 | 0 | 0 | 2 | 32 | 62 | 79 | 53 | 228 |
| Subtotal | 236 | 1322 | 3945 | 6301 | 6085 | 1149 | 564 | 177 | 19779 |

Distribution of discrepancy between predicted level and actual HKDSE level

The discrepancy between predicted level by school and actual HKDSE level derived from the above tables are further summarised in Table 5.

Table 5. Discrepancy between Predicted Level by School and Actual HKDSE Level for CEML

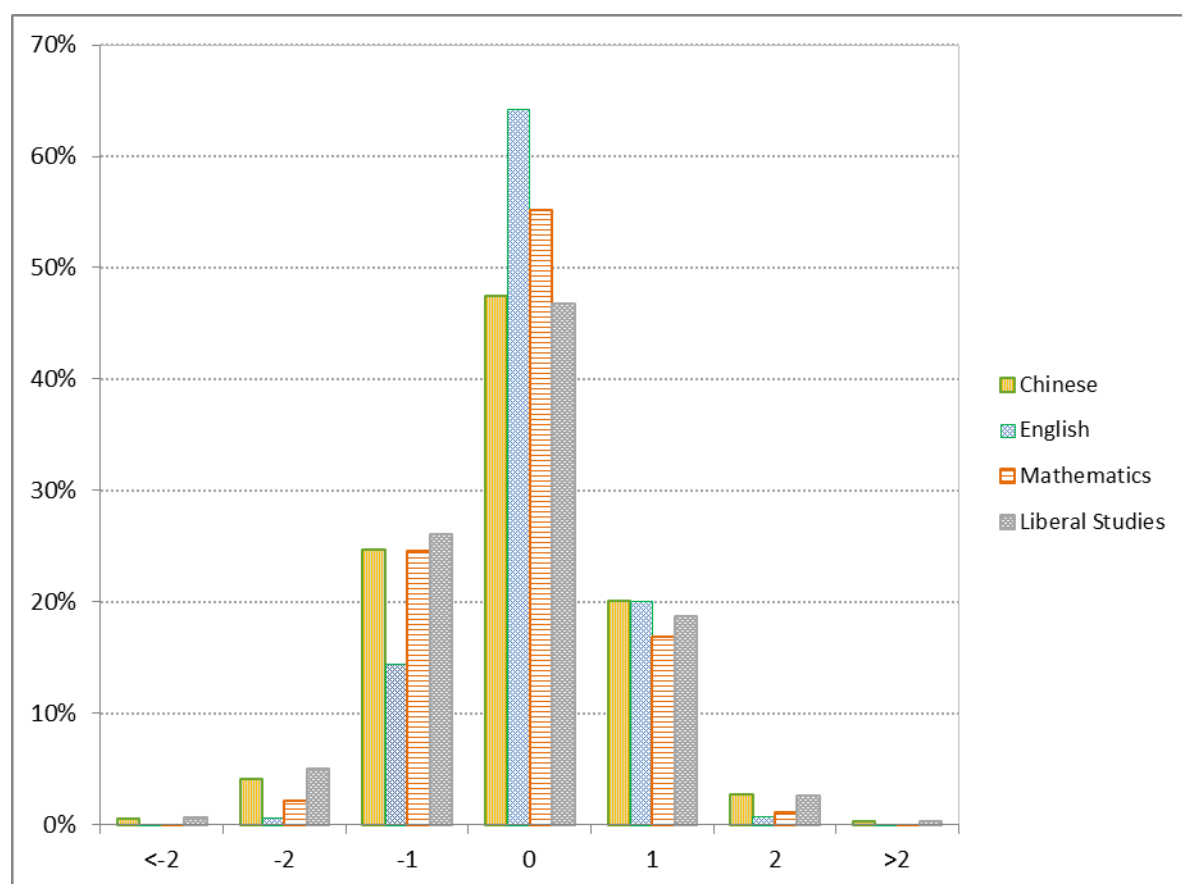
| Subject | | Discrepancy (Predicted Level - Actual Level) | | | | | | |
|-----------------|--------------|--|------|---------|---------------|-------|------|------|
| | | Underestimated | | Perfect | Overestimated | | | |
| | | <-2 | -2 | -1 | 0 | 1 | 2 | >2 |
| Chinese | Count | 112 | 808 | 4857 | 9314 | 3951 | 538 | 64 |
| | % | 0.6% | 4.1% | 24.7% | 47.4% | 20.1% | 2.7% | 0.3% |
| English | Count | 9 | 119 | 2845 | 12639 | 3946 | 144 | 1 |
| | % | 0.0% | 0.6% | 14.4% | 64.1% | 20.0% | 0.7% | 0.0% |
| Mathematics | Count | 15 | 425 | 4845 | 10864 | 3325 | 214 | 11 |
| | % | 0.1% | 2.2% | 24.6% | 55.2% | 16.9% | 1.1% | 0.1% |
| Liberal Studies | Count | 129 | 994 | 5149 | 9247 | 3692 | 508 | 60 |
| | % | 0.7% | 5.0% | 26.0% | 46.8% | 18.7% | 2.6% | 0.3% |
| Total | Count | 265 | 2346 | 17696 | 42064 | 14914 | 1404 | 136 |
| | % | 0.3% | 3.0% | 22.4% | 53.4% | 18.9% | 1.8% | 0.2% |

Broadly speaking, the predicted levels were similar to the actual levels for all four subjects. For predicted level 1 to level 4, most of the actual levels fall within the diagonal cells (i.e., no discrepancy between the predicted and the actual levels) in Tables 1 to 4. This agreement was obvious especially in English Language (with

64.1% correct prediction) and Mathematics (with 55.2% correct prediction).

As shown in Table 5, over half of the cases showed correct prediction. For about 95% of the cases the predicted levels deviated within one level from the actual levels. In general, schools tended to underestimate the levels obtained by their students, especially for Chinese Language and Liberal Studies. The distribution of discrepancy between predicted level and actual HKDSE level for CEML is shown in Figure 1 below.

Figure 1. Distribution of Discrepancy between Predicted Level and Actual HKDSE Level for CEML



Kappa Statistics

Cohen's kappa statistics, which measure the agreement between the predicted levels and actual levels were also considered. Kappa is defined as

$$K = \frac{\sum_d - \sum_{ex}}{N - \sum_{ex}}$$

where,

Σ_d is the total number of agreements obtained by counting the diagonal cells;
 Σ_{ex} is the sum of the expected frequencies for diagonal cell obtained by computing (row total \times column total) / overall total.

Kappa statistics usually have values ranging from 0.0 to 1.0, with a greater value representing a higher degree of agreement*.

Table 6. Kappa Statistics between Predicted Levels by School and Actual HKDSE Levels for CEML

| Subject | Kappa Statistics |
|-----------------|-------------------------|
| Chinese | 0.324 ($p < 0.001$) |
| English | 0.553 ($p < 0.001$) |
| Mathematics | 0.449 ($p < 0.001$) |
| Liberal Studies | 0.310 ($p < 0.001$) |

The predicted and the actual levels were in good agreement in English Language and Mathematics, while fair agreement was observed between the predicted and the actual levels of Chinese Language and Liberal Studies.

Regression Analysis

The relationship between the predicted and the actual levels is shown by the following regression analysis results**:

- Actual HKDSE CHI level = 0.983*(Predicted level of CHI) r = 0.763
- Actual HKDSE ENG level = 0.973*(Predicted level of ENG) r = 0.903
- Actual HKDSE MATH level = 1.005*(Predicted level of MATH) r = 0.881
- Actual HKDSE LS level = 0.987*(Predicted level of LS) r = 0.749

Note:

**Interpretation of Kappa Agreement*

- < 0.01 *No agreement*
- 0.01–0.20 *Low agreement*
- 0.21–0.40 *Fair agreement*
- 0.41–0.60 *Good agreement*
- >0.60 *High agreement*

***The correlation coefficients above should be read with the understanding that the variables (levels) are not continuous variables but ordinal variables.*

Longitudinal Analysis

A longitudinal analysis was conducted to see whether improved familiarity with the HKDSE standards had resulted in more accurate predictions across years. From 2014 to 2017, 59 schools had joined this study for four consecutive years. Table 7 shows the number of candidates in these four years.

Table 7. No. of candidates from 59 schools who joined the study from 2014 to 2017

| Year | No. of Candidates |
|--------------|-------------------|
| 2014 | 8910 |
| 2015 | 8272 |
| 2016 | 7770 |
| 2017 | 7069 |
| Total | 32,021 |

Table 8. Cross-year Comparison of Discrepancy between Predicted Level by School and Actual HKDSE Level for CEML

| | <-2 | -2 | -1 | 0 | 1 | 2 | >2 |
|------------------------|------|------|-------|-------|-------|------|------|
| Chinese | | | | | | | |
| 2014 | 0.6% | 4.5% | 22.9% | 43.8% | 23.8% | 4.0% | 0.4% |
| 2015 | 0.3% | 3.2% | 20.7% | 47.2% | 23.3% | 4.8% | 0.5% |
| 2016 | 0.6% | 5.4% | 27.3% | 45.5% | 18.1% | 2.7% | 0.3% |
| 2017 | 0.7% | 4.9% | 26.0% | 46.3% | 19.2% | 2.8% | 0.3% |
| English | | | | | | | |
| 2014 | 0.0% | 1.8% | 23.4% | 58.9% | 15.0% | 0.8% | 0.0% |
| 2015 | 0.2% | 1.2% | 17.8% | 63.3% | 17.1% | 0.4% | 0.0% |
| 2016 | 0.2% | 1.1% | 24.7% | 62.6% | 10.8% | 0.5% | 0.0% |
| 2017 | 0.1% | 0.9% | 16.5% | 66.0% | 16.0% | 0.5% | 0.0% |
| Mathematics | | | | | | | |
| 2014 | 0.1% | 2.2% | 21.6% | 55.1% | 19.6% | 1.3% | 0.0% |
| 2015 | 0.1% | 2.8% | 25.1% | 53.9% | 16.9% | 1.1% | 0.1% |
| 2016 | 0.0% | 1.8% | 23.7% | 56.2% | 17.0% | 1.0% | 0.2% |
| 2017 | 0.1% | 2.5% | 26.4% | 53.5% | 16.6% | 0.8% | 0.0% |
| Liberal Studies | | | | | | | |
| 2014 | 0.8% | 5.4% | 24.5% | 45.9% | 20.2% | 2.9% | 0.2% |
| 2015 | 0.8% | 5.7% | 25.5% | 44.6% | 20.6% | 2.7% | 0.2% |
| 2016 | 0.6% | 4.8% | 25.8% | 45.9% | 19.4% | 3.1% | 0.4% |
| 2017 | 0.8% | 4.9% | 25.7% | 46.0% | 19.1% | 3.2% | 0.4% |

As shown in Table 8, nearly half of the cases had correct prediction for Chinese Language and Liberal Studies. The predictions of English Language and Mathematics were slightly better and over 53% of the cases showed perfect prediction. It was observed that schools tended to underestimate rather than overestimate their students' attainment for English Language, Mathematics and Liberal Studies, whereas for Chinese Language, the percentages of overestimation were higher than those of underestimation in 2014 and 2015 and vice versa in 2016 and 2017.

The Kappa statistics between predicted levels by school and actual HKDSE level were in moderate agreement for English Language and Mathematics, and fair agreements were observed in Chinese Language and Liberal Studies as shown in Table 9 below.

Table 9. Kappa Statistics between Predicted Level by School and Actual HKDSE Level for CEML

| Subject | 2014 | 2015 | 2016 | 2017 |
|-----------------|-------------|-------------|-------------|-------------|
| Chinese | 0.282 | 0.321 | 0.297 | 0.304 |
| English | 0.488 | 0.539 | 0.530 | 0.575 |
| Mathematics | 0.442 | 0.429 | 0.455 | 0.423 |
| Liberal Studies | 0.290 | 0.280 | 0.294 | 0.289 |

Although it was expected that schools would have more reliable level predictions as teachers became more au fait with the standards and requirements of the HKDSE, it can be concluded from the above results that no consistent increasing trend can be observed in the percentage of correct predictions or the Kappa Statistics.

Further Studies

Data have been collected for the 2018 study which is underway. It is recommended that further investigations be conducted to study the reasons for the consistency or lack thereof in the patterns of school predictions. More importantly, the underlying factors affecting teachers' predictions should be identified in follow-up studies. One possible reason for big discrepancies in prediction is teachers' lack of understanding of the published set of standards, which may create problems in setting teaching goals and thus adversely affect the learning processes. Qualitative studies involving interviews and focus group discussion with teachers may shed light on the major factors affecting the accuracy of schools' predictions.

Acknowledgement

The authors would like to thank colleagues from the Assessment Technology and Research Division and the Examination Systems Unit of the Hong Kong Examinations and Assessment Authority without whose assistance the study could not have been conducted.

References

Education Commission, Reform Proposals for the Education System in Hong Kong, Available from https://www.e-c.edu.hk/doc/en/publications_and_related_documents/education_reform/Edu-reform-eng.pdf. [2000].

HKDSE Assessment Frameworks, Available from http://www.hkeaa.edu.hk/en/hkdse/assessment/subject_information/category_a_subjects/

HKDSE Grading Procedures and Standards-referenced Reporting Booklet, Available from http://www.hkeaa.edu.hk/DocLibrary/Media/Leaflets/HKDSE_SRR_A4booklet_Mar2018.pdf