Opportunities and Challenges in Assessment in the Digital Era

2020.11.12

PROCEEDINGS
Table of Content  目錄

Forum Programme 論壇程序表 ............................................................................................................. 1

Keynote Address 1 主題演講（一） ........................................................................................................ 7
  Integrating Measurement Principles into Formative Assessment
  Dr. Randy E. BENNETT (ETS, Princeton)

Keynote Address 2 主題演講（二） ..................................................................................................... 8
  Urgent CD-CAT Research During The COVID-19 Pandemic
  Prof. CHANG Hua-hua (Purdue University)

Keynote Address 3 主題演講（三） ..................................................................................................... 9
  Assessing Digital Competence
  Prof. Nancy LAW (The University of Hong Kong)

Keynote Address 4 主題演講（四） ................................................................................................... 10
  Making Use of High-Stakes Examinations to Provide Diagnostic Feedback to
  Students and Teachers
  Dr. HUANG Xiaoting (Hong Kong Examinations and Assessment Authority)

Parallel Session 1 分題研討（一）Using Assessment Technology to Inform Learning
  Using Small Data and Big Data to Improve Your Teaching: Essays, Discussions, and
  Class Activities
  Prof. CHIU Ming-ming (The Education University of Hong Kong) ................................................. 11

  Leveraging Modern Psychometrics and Technology to Facilitate Assessment and Learning
  Prof. Jimmy DE LA TORRE (The University of Hong Kong) ............................................................ 12

  Using Formative Assessments to Monitor and Improve Student Learning:
  An Evidence-based Approach
  Dr. Emily OON Pey-tee (University of Macau) .............................................................................. 13

  Applying Automated Scoring Technique to English Writing Tasks: A Preliminary Study
  Dr. JIN Kuan-yu (Hong Kong Examinations and Assessment Authority) ................................. 14
Parallel Session 2  分題研討（二）Using Assessment Data to Facilitate Learning in School Context

Uptake of Student Self-assessment in Schools
Dr. YAN Zi (The Education University of Hong Kong) ................................................................. 15

有效利用電子評核系統進行進展性評估
文可為副校長（樂善堂余近卿中學） .................................................................................. 16

An Explorative Study of Candidates’ Performance in the HKDSE Chemistry Examinations
Dr. Raymond FONG Wai-hung (Education Bureau)
Dr. LI Tak-man (Hong Kong Examinations and Assessment Authority) ........................................ 17

在學校進行電子評卷與試後數據分析
蕭偉樂博士（香港考試及評核局） ....................................................................................... 19

Parallel Session 3  分題研討（三）HKDSE-related Studies

A Feasibility Study on a Computer-based English Language Speaking Test
Dr. Cameron SMART (Hong Kong Examinations and Assessment Authority) ............................. 20

A Study on the Contribution of Fitness Factors to the Performance of Physical Activities of the HKDSE Physical Education
Mr. Kevin KAM (The Education University of Hong Kong)
Dr. Steven YIP (Hong Kong Examinations and Assessment Authority) ........................................ 21

A Study of the Performance Difference in BAFS Paper 1 between Candidates Taking the Accounting Module and the Business Management Module
Ms. Sue SO (Hong Kong Examinations and Assessment Authority) .............................................. 23

中國語文試卷一增設甲部指定文言篇章是否有助中文學習
葉子彬先生（香港考試及評核局） .......................................................................................... 24

Parallel Session 4  分題研討（四）Applications of CBT in Large Scale High-Stakes Examinations

大規模計算機化考試在全國大學英語四、六級口語考試中的應用
趙英華副處長（教育部考試中心） ....................................................................................... 25

計算機輔助測試在廣東省高考英語聽說中的實踐和探索
范韶彬副院長（廣東省教育考試院） .................................................................................. 26

AI 在大規模高利害考試中的研究與應用
朱汝光處長（北京教育考試院） .......................................................................................... 27

雲網融合的在線考試在省級大規模考試中的探索
范鵬副研究員（天津市教育招生考試院） ............................................................................ 28
**Forum Programme (Morning Sessions)**

**Venue:** 4/F Lecture Theatre, EDB Kowloon Tong Education Services Centre (West Block)

**Programme**

<table>
<thead>
<tr>
<th>Hong Kong Time (HKT)</th>
<th>Programme</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:45am – 9:15am</td>
<td>Registration</td>
</tr>
<tr>
<td>9:15am – 9:30am</td>
<td>Opening Ceremony</td>
</tr>
<tr>
<td></td>
<td>Welcome Address</td>
</tr>
<tr>
<td></td>
<td>by Mr. Samuel YUNG Wing-ki, SBS, MH, JP</td>
</tr>
<tr>
<td></td>
<td>(Chairman, Hong Kong Examinations and Assessment Authority)</td>
</tr>
<tr>
<td></td>
<td>Introductory Address</td>
</tr>
<tr>
<td></td>
<td>by Ir. Dr. SO Kwok-sang</td>
</tr>
<tr>
<td></td>
<td>(Secretary General, Hong Kong Examinations and Assessment Authority)</td>
</tr>
<tr>
<td>9:30am – 10:00am</td>
<td>Keynote Address 1</td>
</tr>
<tr>
<td></td>
<td><strong>Integrating Measurement Principles into Formative Assessment</strong></td>
</tr>
<tr>
<td></td>
<td>Dr. Randy E. BENNETT * (ETS, Princeton)</td>
</tr>
<tr>
<td></td>
<td>(English 英語)</td>
</tr>
<tr>
<td>10:00am – 10:30am</td>
<td>Keynote Address 2</td>
</tr>
<tr>
<td></td>
<td><strong>Urgent CD-CAT Research During The COVID-19 Pandemic</strong></td>
</tr>
<tr>
<td></td>
<td>Prof. CHANG Hua-hua * (Purdue University)</td>
</tr>
<tr>
<td></td>
<td>張華華教授 (美國普渡大學)</td>
</tr>
<tr>
<td></td>
<td>(English 英語)</td>
</tr>
<tr>
<td>10:30am – 10:45am</td>
<td>Break</td>
</tr>
<tr>
<td>10:45am – 11:15am</td>
<td>Keynote Address 3</td>
</tr>
<tr>
<td></td>
<td><strong>Assessing Digital Competence</strong></td>
</tr>
<tr>
<td></td>
<td>Prof. Nancy LAW (The University of Hong Kong)</td>
</tr>
<tr>
<td></td>
<td>羅陸慧英教授 (香港大學)</td>
</tr>
<tr>
<td></td>
<td>(English 英語)</td>
</tr>
<tr>
<td>11:15am – 11:45am</td>
<td>Keynote Address 4</td>
</tr>
<tr>
<td></td>
<td><strong>Making Use of High-Stakes Examinations to Provide Diagnostic Feedback to Students and Teachers</strong></td>
</tr>
<tr>
<td></td>
<td>Dr. HUANG Xiaoting (Hong Kong Examinations and Assessment Authority)</td>
</tr>
<tr>
<td></td>
<td>黃曉婷博士 (香港考試及評核局)</td>
</tr>
<tr>
<td></td>
<td>(English 英語)</td>
</tr>
<tr>
<td>11:45am – 12:00nn</td>
<td>Concluding Remarks</td>
</tr>
<tr>
<td></td>
<td>Concluding Remarks by Ir. Dr. SO Kwok-sang</td>
</tr>
<tr>
<td></td>
<td>(Secretary General, Hong Kong Examinations and Assessment Authority)</td>
</tr>
<tr>
<td></td>
<td>總結：蘇國生博士 (香港考試及評核局秘書長)</td>
</tr>
</tbody>
</table>

* *Speakers will conduct the presentation through webcast.* 講者將透過網上視像形式發表演講。
Parallel Session 1 分題研討（一）  
Using Assessment Technology to Inform Learning

Venue: WP01, EDB Kowloon Tong Education Services Centre (West Block)  
地點： 教育局九龍塘教育服務中心（西座）WP01 室

<table>
<thead>
<tr>
<th>Hong Kong Time (HKT)</th>
<th>Programme</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:30pm – 2:00pm</td>
<td>Registration</td>
</tr>
</tbody>
</table>
| 2:00pm – 2:40pm      | * Using Small Data and Big Data to Improve Your Teaching: Essays, Discussions, and Class Activities  
                        Prof. CHIU Ming-ming (The Education University of Hong Kong)  
                        趙明明教授（香港教育大學）  
                        (English 英語) |
| 2:40pm – 3:20pm      | * Leveraging Modern Psychometrics and Technology to Facilitate Assessment and Learning  
                        Prof. Jimmy DE LA TORRE * (The University of Hong Kong)  
                        (English 英語) |
| 3:20pm – 4:00pm      | * Using Formative Assessments to Monitor and Improve Student Learning: An Evidence-based Approach  
                        Dr. Emily OON Pey-tee * (University of Macau)  
                        溫佩娣博士（澳門大學）  
                        (English 英語) |
| 4:00pm – 4:40pm      | * Applying Automated Scoring Technique to English Writing Tasks: A Preliminary Study  
                        Dr. JIN Kuan-yu (Hong Kong Examinations and Assessment Authority)  
                        金冠宇博士（香港考試及評核局）  
                        (English 英語) |

* Speakers will conduct the presentation through webcast. 講者將透過網上視像形式發表演講。
Parallel Session 2 分題研討（二）
Using Assessment Data to Facilitate Learning in School Context

Venue: 4/F Lecture Theatre, EDB Kowloon Tong Education Services Centre (West Block)
地點：教育局九龍塘教育服務中心（西座）四樓演講廳

<table>
<thead>
<tr>
<th>Hong Kong Time (HKT)</th>
<th>Programme 程序</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:30pm – 2:00pm</td>
<td>Registration 登記入座</td>
</tr>
</tbody>
</table>
| 2:00pm – 2:40pm      | Uptake of Student Self-assessment in Schools
                      | Dr. YAN Zi (The Education University of Hong Kong)
                      | 晏子博士（香港教育大學）
                      | (Cantonese 廣東話) |
| 2:40pm – 3:20pm      | 有效利用電子評核系統進行進展性評估
                      | Mr. MAN Ho-wai (LST Yu Kan Hing Secondary School)
                      | 文可為副校長（樂善堂余近卿中學）
                      | (Cantonese 廣東話) |
| 3:20pm – 4:00pm      | An Explorative Study of Candidates’ Performance in the HKDSE Chemistry Examinations
                      | Dr. Raymond FONG Wai-hung (Education Bureau)
                      | 方偉雄博士（教育局）
                      | Dr. LI Tak-man (Hong Kong Examinations and Assessment Authority)
                      | 李德文博士（香港考試及評核局）
                      | (Cantonese 廣東話) |
| 4:00pm – 4:40pm      | 在學校進行電子評卷與試後數據分析
                      | Dr. SIU Wai-lok (Hong Kong Examinations and Assessment Authority)
                      | 蕭偉樂博士（香港考試及評核局）
<pre><code>                  | (Cantonese 廣東話) |
</code></pre>
<table>
<thead>
<tr>
<th>Hong Kong Time (HKT)</th>
<th>Programme</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:30pm – 2:00pm</td>
<td>Registration</td>
</tr>
</tbody>
</table>
| 2:00pm – 2:40pm     | *A Feasibility Study on a Computer-based English Language Speaking Test*  
Dr. Cameron SMART (Hong Kong Examinations and Assessment Authority)  
(English 英語) |
| 2:40pm – 3:20pm     | *A Study on the Contribution of Fitness Factors to the Performance of Physical Activities of the HKDSE Physical Education*  
Mr. Kevin KAM (The Education University of Hong Kong)  
甘偉強先生 (香港教育大學)  
Dr. Steven YIP (Hong Kong Examinations and Assessment Authority)  
葉弘博士 (香港考試及評核局)  
(Cantonese 廣東話) |
| 3:20pm – 4:00pm     | *A Study of the Performance Difference in BAFS Paper 1 between Candidates Taking the Accounting Module and the Business Management Module*  
Ms. Sue SO (Hong Kong Examinations and Assessment Authority)  
蘇慧君女士 (香港考試及評核局)  
(Cantonese 廣東話) |
| 4:00pm – 4:40pm     | *中國語文試卷一增設甲部指定文言篇章是否有助中文學習*  
Mr. YIP Chi-bun (Hong Kong Examinations and Assessment Authority)  
葉子彬先生 (香港考試及評核局)  
(Cantonese 廣東話) |
Parallel Session 4  分題研討（四）
Applications of CBT in Large Scale High-Stakes Examinations

Venue:  W134, EDB Kowloon Tong Education Services Centre (West Block)

<table>
<thead>
<tr>
<th>Hong Kong Time (HKT)</th>
<th>Programme</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:30pm – 2:00pm</td>
<td>Registration</td>
</tr>
</tbody>
</table>
| 2:00pm – 2:40pm      | 大規模計算機化考試在全國大學英語四、六級口語考試中的應用  
                        趙英華副處長*（教育部考試中心）  
                        （普通話）                      |
| 2:40pm – 3:20pm      | 計算機輔助考試在廣東省高考英語聽說中的實踐和探索  
                        范韶彬副院長*（廣東省教育考試院）  
                        （普通話）                      |
| 3:20pm – 4:00pm      | AI 在大規模高利害考試中的研究與應用  
                        朱汝光處長*（北京教育考試院）  
                        （普通話）                      |
| 4:00pm – 4:40pm      | 雲網融合的在線考試在省級大規模考試中的探索  
                        范鵬副研究員*（天津市教育招生考試院）  
                        （普通話）                      |

* Speakers will conduct the presentation through webcast. 講者將透過網上視像形式發表演講。
Presentation Abstracts

演講摘要
Keynote Address 1

*Integrating Measurement Principles into Formative Assessment*

Dr. Randy E. BENNETT (ETS, Princeton)

Abstract

This presentation will discuss the application of measurement principles to classroom formative assessment. The presentation applies the perspective of evidentiary reasoning to formative assessment. The presentation will describe four fundamental acts that can serve as a frame for organizing and implementing formative assessment. The acts are: 1) engineering opportunities to observe evidence of the competencies about which we wish to make claims; 2) inferentially connecting that evidence to meaningful characterizations of individuals and groups; 3) acting on those characterizations (including giving feedback, making instructional adjustments and plans); and 4) evaluating the quality and impact of those characterizations and of the decisions based on them.

About the Speaker

Randy Bennett is Norman O. Frederiksen Chair in Assessment Innovation at Educational Testing Service in Princeton, NJ. Since the 1980s, he has conducted research on integrating advances in cognitive science, technology, and measurement to create new approaches to assessment intended to have positive impact on teaching and learning. For his work, he was given the ETS Senior Scientist Award in 1996, the ETS Career Achievement Award in 2005, the Teachers College, Columbia University Distinguished Alumni Award in 2016, Fellowship status in the American Educational Research Association (AERA) in 2017, the National Council on Measurement in Education's (NCME) Bradley Hanson Award for Contributions to Educational Measurement in 2019 (with H. Guo, M. Zhang, and P. Deane), and the E. F. Linquist Award from AERA and ACT in 2020.

Randy Bennett is Past President of the International Association for Educational Assessment (IAEA), which is primarily constituted of governmental and NGO measurement organizations throughout the world. He is also Past President of NCME, whose members are employed in universities, testing organizations, education departments, and school districts.

His Google Knowledge Panel can be found at [https://g.co/kgs/XZRtnF](https://g.co/kgs/XZRtnF)
Keynote Address 2

Urgent CD-CAT Research During The COVID-19 Pandemic

Prof. CHANG Hua-hua (Purdue University)

Abstract

With the COVID-19 pandemic, many schools have canceled in-person classes and moved to online instructions, which has created enormous challenges for both instructors and students. It is likely that an increased number of courses will remain online or hybrid even after the pandemic ends. In today’s presentation we will discuss the possibility of utilizing Cognitive Diagnostic Computerized Adaptive Testing (CD-CAT) to support everyday instructions whether in-class, online or hybrid. CAT is rapidly transforming testing from unaccommodating ranking measures into flexible and informative tools that can be used to address the compelling needs of various stakeholders in education. CD-CAT is expected to provide teachers with precise information about each student’s problem areas so that, even when they cannot see the confusion in students’ faces, they will know when they need to step in with clarifying information. A build-in time-management mechanism will greatly enhance test-security and make students take in-class exam at home possible.

About the Speaker

Dr. Chang Hua-hua is the Charles R. Hicks Chair Professor in the Department of Educational Studies at Purdue University. Dr. Chang’s interests are broad, encompassing both theoretical development and applied methodologies in psychometrics, including computerized adaptive testing, cognitive diagnosis, asymptotic properties in IRT, and test equity. Most recently, his work has been concentrated on developing web-based assessments to facilitate individualized learning. Dr. Chang served Chang Jiang Scholar Chair Professor at East China Normal University (2014-2017), twice served as Fulbright Specialist (Colombia 2019 and Australia 2005). He is a Past-President of the Psychometric Society and a Fellow of both the American Educational Research Association and American Statistical Association. Dr. Chang is the recipient of the 2017 AERA’s E. F. Lindquist Award.
Keynote Address 3

Assessing Digital Competence

Prof. Nancy LAW (The University of Hong Kong)

Abstract

A lot of changes have taken place in terms of both curriculum and pedagogy over the past two decades, in Hong Kong and in other education systems. An important driving force for such changes is the need for education to foster 21st century competences. There have also been some changes in assessment practices to emphasize understanding, communication and argumentation within existing school subjects. On the other hand, digital competence as a generic 21st century skill is overlooked in current curriculum and assessment provisions and practices. International and local assessment research have shown serious inadequacies in Hong Kong students’ digital competence, which additionally threatens students’ ability to learn effectively through online means. In this paper, I will introduce current research on curriculum and assessment of digital competence, and argue for strong policy support in Hong Kong to develop an assessment system for digital competence to support learning and teaching for this very important 21st century capacity.

About the Speaker

Nancy Law is a professor in the Faculty of Education at the University of Hong Kong and the Founding Director for the Centre for Information Technology in Education. She is a Fellow of the International Society of the Learning Sciences. Her research focuses on designing and leading scalable technology-enhanced learning-innovations, and investigations of students’ development as digital citizens from childhood to early adulthood. She has served on a number of policy advisory boards/working groups related to ICT in education for the University of Hong Kong, the Hong Kong SAR government and other community, government groups and institutions. She has also been contributing as expert consultant to the European Commission, UNESCO and OECD on various aspects of technology-enhanced learning.
Abstract

Making use of the past HKDSE papers, the HKDSE Diagnostic Feedback System (DFS) aims at providing users feedback that can facilitate learning and instruction. Specifically, students who take a set of questions online can receive instant reports that pinpoint specific areas of concern in their study. Meanwhile, teachers receive detailed item analysis reports as well as reports on their students’ performance so that they can make necessary adjustments to their teaching plan. The presentation demonstrates how the system works, and briefly explains the quantitative and qualitative work involved in developing the system. Future development plan is also touched upon.

About the Speaker

Dr. Huang Xiaoting is the Director – Examinations, Assessment and Research of the Hong Kong Examinations and Assessment Authority. She is in charge of formulating strategies and policies governing research initiatives, promoting assessment for learning, conducting analytics in support of the standards of the HKDSE. Prior to joining the HKEAA, she was a faculty at Peking University. Her research area spans a range of issues from test reliability and validity, item response modelling, to policy issues on the use of assessment data.
Using Small Data and Big Data to Improve Your Teaching: Essays, Discussions, and Class Activities

Prof. CHIU Ming-ming (The Education University of Hong Kong)

Abstract

Computers and videos are creating bigger, complex data sets (homework, tests, essays, online forums, classroom activities). By combining computers, statistics, mathematics, and linguistics to analyze these big data (analytics), we can better understand our students’ strengths and weaknesses to inform our teaching. Illustrative cases include analytics of students’ (a) essays for content and creativity; (b) online discussions for unproductive sequences to alert teachers; and (c) classroom activities for attention and engagement. Parallel to big data analytics, I show how a teacher can do simple, informative analyses of small data (e.g., students in one class).

About the Speaker

Chiu Ming-ming is Chair Professor of Analytics and Diversity in the Special Education and Counseling department and Director of the Assessment Research Center at The Education University of Hong Kong. A graduate of Columbia (BS, computer science), Harvard (EdM, interactive technology) and UC-Berkeley (PhD, education), he is Analytics Advisor to South Korea's Minister of Corporations and advises Qatar’s Ministry of Education and China’s Ministry of Education. He invented (a) statistical discourse analysis to model online and face-to-face conversations (one of the best 50 learning science ideas –International Society of the Learning Sciences), (b) multilevel diffusion analysis to detect corruption in the music industry and how ideas/behaviors spread through populations, (c) artificial intelligence Statistician, and (d) online detection of sexual predators. His 55 grants (HK$82 million) yielded 224 publications (137 journal articles; 8,000+ citations), 13 keynote speeches, 4 television broadcasts, 17 radio broadcasts, and 165 news articles in 21 countries. He creates and applies automatic statistical analyses to Big Data.
**Abstract**

Many educational researchers and practitioners are interested in using educational assessment to improve student learning. However, as two distinct components, assessment and learning need to be integrated before the former can be used to inform the latter. In this presentation, I will discuss cognitive diagnosis modeling as a coherent framework for integrating assessment and learning. Specifically, I will introduce cognitive diagnosis models (CDMs), discuss their unique features, and highlight how they differ from traditional psychometric models. In addition to assessment, instructional materials based on the same framework are needed to facilitate learning. By leveraging technology, computerized adaptive testing and ancillary information can be used to further capitalize on the advantages of CDMs and make diagnostic testing more efficient. Similarly, technology can also be leveraged to determine the extent to which different instructional materials can be tailored to optimize learning. The presentation will conclude with a discussion of some of the challenges, recent developments, and possible future directions in the area.

**About the Speaker**

Jimmy de la Torre is a Professor in the Human Communication, Development, and Information Sciences Unit of the Faculty of Education at The University of Hong Kong. He is also currently a Chair Professor at the National Taichung University of Education in Taiwan, and an Honorary Professor at the Universidad Autonoma de Madrid in Spain. His primary research interests are in the field of educational and psychological measurement, with a particular emphasis on item response theory, cognitive diagnosis modeling, and the use of assessment to inform instruction and learning. In 2009, he was named by the White House as one of the recipients of the Presidential Early Career Awards for Scientists and Engineers. He also received the Jason Millman Promising Measurement Scholar Award in 2009 from the National Council on Measurement in Education. He was editor-in-chief of the Journal of Educational Measurement, and is currently an associate editor of Applied Psychological Measurement and Frontiers in Education. For over 15 years now, Jimmy has conducted more than 40 training and professional development workshops on cognitive diagnosis modeling in more than a dozen countries and four continents.
Using Assessment Technology to Inform Learning

Using Formative Assessments to Monitor and Improve Student Learning: An Evidence-based Approach

Dr. Emily OON Pey-tee (University of Macau)

Abstract

Bloom’s cognitive taxonomy provides a way to classify the level of knowledge required by students to meet specific learning objectives (Bloom, 1956). As such, it can be used as a measurement tool that helps teachers understand cognitive ability of students with respect to intended learning outcomes (Krathwohl, 2002) through the use of classroom formative assessments. Formative assessments provide feedback that can be used to improve students learning progressions and help teachers to adjust their teaching to better meet student and learning needs. Often, teachers are unsure of how formative assessment can be designed and carried out so to obtain information that provides meaningfully information on students understanding. Following Wilson (2004)’s four building blocks, which are (a) construct map, (b) item design, (c) outcome space, and (d) measurement model, we have developed and piloted a test instrument consisting of 21 multiple choice questions that measure arithmetic understanding learning outcomes of primary three students in Macau. In this presentation we will discuss how to develop good formative assessments, how the information obtained can be used to improve teaching and learning and the importance of monitoring and using student learning progressions as a means to improve student outcomes in STEM courses.

About the Speaker

Emily Oon is an associate professor in the Faculty of Education at the University of Macau. Her main research interests include psychometric assessments of STEM high stakes survey and test items, (2) development of Modern K-12 integrated STEM education in Macau, and (3) understanding and improving K-12 classroom assessment. She also helps to review papers submitted to international journals, including (1) Science & Education, (2) Asia-Pacific Forum on Science Learning and Teaching, (3) The Asia Pacific Education Researcher, (4) Studies in Higher Education, (5) Studies in Educational Evaluation, (6) Assessment and Evaluation in Higher Education, etc.
**Applying Automated Scoring Technique to English Writing Tasks: A Preliminary Study**

Dr. JIN Kuan-yu (Hong Kong Examinations and Assessment Authority)

**Abstract**

Since the beginning of 21st century, automated scoring technique (AST) has been developed and applied to score English essays. In this study, we collaborated with iFlyTek to explore the feasibility of applying AST to Hong Kong Diploma of Secondary Education (HKDSE) English Language examination writing tasks. The preliminary results showed that iFlyTek’s engine yielded similar scores to those given by human raters in the HKDSE tasks. Limitations of AST to existing examinations hosted by the Authority and possible further studies are also discussed.

**About the Speaker**

Dr. Jin Kuan-yu is a manager in the Assessment Technology and Research Division at the Hong Kong Examinations and Assessment Authority. He received his PhD from the Education University of Hong Kong in 2017. His primary research interests include Rasch measurement, item response theory models, and psychometrics.
Abstract

Self-assessment is a core skill in developing students’ self-regulated learning and life-long learning. Past research has shown a consistently positive correlation between student self-assessment and learning outcomes. This presentation discusses conditions that facilitate the uptake of student self-assessment in schools. It presents one self-assessment diary intervention embedded in daily teaching as an example and examines its effectiveness on enhancing students’ academic achievement, self-regulation, and motivation on secondary one students. The principles in designing sustainable self-assessment interventions are discussed.

About the Speaker

Dr. YAN Zi is Associate Professor at the Department of Curriculum and Instruction, the Education University of Hong Kong. He is also the Co-Director of the Centre for Excellence in Learning and Teaching (CELT) and Associate Director of the Assessment Research Centre (ARC). Dr. Yan obtained his PhD from James Cook University, Australia. His research interests focus on the use of Rasch measurement for assessment in school and higher education contexts, with an emphasis on student self-assessment and self-regulated learning.
有效利用電子評核系統進行進展性評估

文可為副校長（樂善堂余近卿中學）

摘要

隨著無線網絡環境優化及流動學教裝置的普及，加上雲端平台工具的應用，推行電子評核於學教日趨方便及具成效。通過電子評核系統能夠有效收集、儲存及分析評估數據，老師進行專業反思從而優化學教策略。

是次分享重點講解學界主流電子評核系統及工具進的使用，如何有效進行進展性評估，及就學校的應用案例，分享推展有關計劃時需要關注及如何規劃相關策略。

講者簡介

文可為副校長對推動學校資訊科技教育不遺餘力。曾獲行政長官卓越教學獎，並為教育局資訊科技教育卓越中心擔任專業發展課程導師。
**Parallel Session 2 分題研討（二）**

*Using Assessment Data to Facilitate Learning in School Context*

**An Explorative Study of Candidates’ Performance in the HKDSE Chemistry Examinations**

Dr. Raymond FONG Wai-hung (Education Bureau)

Dr. LI Tak-man (Hong Kong Examinations and Assessment Authority)

Abstract

**Objectives:**
The study is intended to explore candidates’ performance in relation to attributes of question, i.e. “depth of knowledge”, “number of knowledge points / steps”, “unfamiliarity” and “cross-topic” (abbreviated as DOK31, NKP/S, UF and CT respectively). Furthermore, the study also aims to provide qualitative observations on candidates’ performance in selected samples of questions.

**Literature review:**
To develop a better background knowledge of the study, research articles related to chemistry knowledge triplet, chemistry textbooks, performance of candidates in examinations, cognitive demand / level of difficulty of assessment tasks, etc. were scrutinised.

**Scope:**
Three different question types (i.e. multiple-choice question (MCQ), short/structured question (SQ) and paragraph-length question (PQ)) from HKDSE2 2012 to HKDSE 2018 were included in the study. To provide more meaningful findings, only questions from three topics, “Redox reactions, chemical cells and electrolysis”, “Rate of reaction” and “Industrial chemistry” were included.

**Quantitative findings:**
1. Descriptive statistics of performance in different question types:

   ![Figure 1](image1.png)

   (Legend: **mean performance, **= 50 percentile, **= interquartile range, **= maximum / minimum, **= outliers)

<table>
<thead>
<tr>
<th>Performance of MCQ against DOK3:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Perform (%)</strong></td>
</tr>
<tr>
<td><strong>DOK3</strong> 1 (N=11) <strong>2 (N=18) 3 (N=6)</strong></td>
</tr>
<tr>
<td>g</td>
</tr>
<tr>
<td>σ</td>
</tr>
<tr>
<td>Range</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Performance of SQ against DOK3:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Perform (%)</strong></td>
</tr>
<tr>
<td><strong>DOK3</strong> 1 (N=18) <strong>2 (N=24) 3 (N=27)</strong></td>
</tr>
<tr>
<td>g</td>
</tr>
<tr>
<td>σ</td>
</tr>
<tr>
<td>Range</td>
</tr>
</tbody>
</table>

**1** DOK3 refers to Webb’s Depth of Knowledge Level 1, 2 and 3; with Level 4 (or Extended Thinking) excluded.

**2** HKDSE is the abbreviation of Hong Kong Diploma of Secondary Education.

**3** Only questions related to rate of reaction from the topic “Industrial chemistry” were included.

**4** Selected examples of findings only, more findings will be discussed.
Qualitative observations:

3. Some questions with exceptionally high or low performance were analysed qualitatively:

<table>
<thead>
<tr>
<th>Question</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>The following equation shows the reaction when a secondary cell is discharging: 2N₂O₄(0) + Cd(0) → 2N₂(OH)₂(0) + Cd(OH)₂(0) Which of the following half equations shows the change at the negative electrode when the cell is being recharged?</td>
<td></td>
</tr>
<tr>
<td>A. Cd(s) + 2OH⁻(aq) → Cd(OH)₂(s) + 2e⁻</td>
<td></td>
</tr>
<tr>
<td>B. Cd(OH)₂(s) + 2e⁻ → Cd(s) + 2OH⁻(aq)</td>
<td></td>
</tr>
<tr>
<td>C. Ni(OH)₂(s) + H₂O(l) → Ni(OH)₃(s) + H₂O(l) + e⁻</td>
<td></td>
</tr>
<tr>
<td>D. Ni(OH)₃(s) + 3OH⁻(aq) → Ni(OH)₂(s) + OH⁻(aq)</td>
<td></td>
</tr>
</tbody>
</table>

The low performance was likely related to candidates’ limited macroscopic related or daily-life experience of using this type of rechargeable secondary cell, and had less than satisfactory understanding on this kind at a submicroscopic level of representation, i.e. using the chemical equations and ionic equations (Johnstone, 1991) to present their answers.

Closing remarks:

Limitations of the study, suggestions for future study and some additional remarks will be discussed.

About the Speaker

Dr. Raymond FONG Wai-hung
Former teacher and subject panel head in secondary school;
Former Senior Curriculum Development Officer of the Education Bureau, overseeing the development of the Chemistry curriculum and relevant teaching resources, and providing relevant professional development courses for teachers and laboratory technicians.

Dr. LI Tak-man
Former teacher and subject panel head in secondary schools;
Former Curriculum Development Officer of the Education Bureau;
Senior Manager of the HKEAA, overseeing the Mathematics and Science subjects and in-charge of the overall administration for the School-based Assessments for all HKDSE subjects.
在學校進行電子評卷與試後數據分析

蕭偉樂博士（香港考試及評核局）

摘要

考評局發展了一套電子評卷系統，方便老師利用平板電腦批改試卷。系統包括兩個相連的部分：作為整體系統和帳戶管理的網站和實際執行評卷操作的應用程式。利用電子模式評卷，不但方便老師儲存各題的分數作日後數據分析之用，更重要是讓老師可以選取學生的佳作或常犯錯誤作為具體例子，對學生的學習作出回饋，實踐評核促進學習。

在試後數據分析方面，考評局的評核質素保證平台(AQP)為學校提供一套功能齊全的網上平台，讓老師自行上載數據，並獲得即時分析結果。分析報告除了檢視試卷整體的信度外，更提供個別題目難度和區分度的指標，可以作為日後改善題目設計的參考。就個別學生成績分析而言，分析報告中的學生題目表(S-P chart)、題目與受試者能力圖(item-person map)、以及修正注意係數(MCI)等都能為老師提供有用的數據，以識別個別學生下一階段的學習需要。

講者簡介

蕭偉樂博士為香港考試及評核局評核科技及研究部主管。在考評局十多年的工作中，致力推動中小學進行評核後數據分析，協助老師以實際數據支持決策，更有效辨識個別學生的學習需要和改善評核設計形式。
Abstract

This presentation will report on a feasibility study on a computer-based English Language speaking test, which is being conducted by the HKEAA. In the presentation, the different stages of the study will be described, including the devising of a computer-based assessment framework and the conducting of a pilot with a prototype of the test. Statistical and linguistic analyses related to the pilot will also be presented. Implications of the study will be discussed.

About the Speaker

Dr. Smart is an English Language Subject Manager at the HKEAA. He has been working in the field of EFL teaching and testing for the last 32 years, having worked in Scotland, Spain, Hong Kong and England.
**A Study on the Contribution of Fitness Factors to the Performance of Physical Activities of the HKDSE Physical Education**

Mr. Kevin Kam (The Education University of Hong Kong)
Dr. Steven Yip (Hong Kong Examinations and Assessment Authority)

Abstract

There is a myriad of field test protocols for measuring physical fitness of students. HKDSE Physical Education (PE) has been adopting testing protocols similar to those adopted by the School Physical Fitness Award Scheme of the Education Bureau. Sit-ups is used for testing muscular endurance; pull-ups and flexed arm hang for muscular strength; sit and reach for flexibility and 1609m run/walk for cardiovascular endurance. Previous researches have shown that there are positive correlations between physical fitness performance and the performance of physical activities. However, the scope of these researches is restricted to only one or a few physical activities. This research has investigated systematically the relationship between the attainment in physical fitness and corresponding performance in a variety of physical activities of over 5,500 students taking the HKDSE PE examinations from 2012 to 2018. The research results show that there are positive correlations between physical activity performance and physical fitness performance. However, the degree of correlations varies across activities. It is observed that the nature of the sports may influence the degree of correlation. For instance, track events in Athletics have a higher correlation with 1609m run/walk test. Gymnastics have a higher correlation with sit-and-reach test. All these findings are consistent with the idiosyncrasies of the physical activities concerned. The findings of this research provide useful information to students when they choose the optional physical activities in HKDSE PE. They are advised to be aware of their physical fitness level in different components and understand that different physical fitness components may influence their performance in physical activities.

About the Speaker

Kevin Kam is currently a Senior Lecturer and an Associate Head (Learning and Teaching) of the Department of Health and Physical Education, the Education University of Hong Kong. Kevin completed his teacher training in PE and Mathematics from Northcote College of Education. He holds a Bachelor of Education (Honours) degree from the University of Liverpool and a Master of Education degree from the University of Hong Kong. He is also an ACSM Certified Exercise Physiologist. He is currently undertaking his Doctor of Philosophy degree from the University of
Queensland which seeks to explore the knowledge and identity development of beginning PE teachers in Hong Kong from a qualitative research. Before joining the University, Kevin had taught PE and served as Chairman of PE subject panel in a secondary school. He coached a wide range of sports teams in the school. He had taught in tertiary institution for more than 25 years. He taught a wide variety of courses and coached the university sports teams. Kevin now specializes in pedagogy and professional development culture of PE teachers and coaches. He is currently teaching curricular and pedagogical studies in PE, philosophical and historical foundations of PE and sports. He can also teach a wide variety of sports skills, like soccer, golf, physical fitness. His other academic interests include Health and PE Curriculum, Pedagogy, Sports and Society. He has an established track record in developing professional development of PE teachers by researching the use of different teaching and curriculum models, developing a learning community by organizing PE teacher conference, and professional development by action research in physical education. As a PE Teacher Educator and a Coach Educator, he also involved a lots of community services on professional development for PE teachers, curriculum and assessment development in PE, school management, and coach education programme.

Steven Yip is currently the subject manager of PE in the HKEAA. He has been coordinating the assessment development of HKCEE/HKDSE PE in the past 15 years.
A Study of the Performance Difference in BAFS Paper 1 between Candidates Taking the Accounting Module and the Business Management Module

Ms. Sue SO (Hong Kong Examinations and Assessment Authority)

Abstract

During the school consultation about the future development of business subjects conducted in 2017, it was claimed that candidates taking the Accounting Module are normally stronger in numerical skills but weaker in writing, therefore they may have poorer performance in the short questions set in HKDSE BAFS Paper 1, when compared with the performance of candidates taking the Business Management Module. The purpose of this study is to investigate whether the abovementioned claim is supported by past HKDSE BAFS examinations statistics. By conducting a differential item functioning (DIF) analysis to BAFS Paper 1 items from 2015 to 2019, it was revealed that Accounting candidates performed better in some questions while Business Management candidates performed better in others. There is no evidence to show that Accounting candidates have poorer performance in the short questions set in HKDSE BAFS Paper 1, as compared with that of Business Management candidates.

About the Speaker

Ms Sue So is the Subject Manger of the HKDSE Business, Accounting and Financial Studies.
中國語文試卷—增設甲部指定文言篇章是否有助中文學習

葉子彬先生 (香港考試及評核局)

摘要

中國語文卷一自 2018 年始增設甲部考題，考核 12 篇指定文言篇章。增設考核範文的目的是以此作為培養閱讀能力的起點，透過研習經典而能觸類旁通，從而提升其閱讀理解能力，惟增設範文是否可以收到成效仍待觀察。故藉此研究探討增設範文的效用。本研究目的在於初步引證增設指定篇章是否可以提升考生閱讀課外白話文或文言文的能力，甚或整體的閱讀能力。同時，藉此反思考核指定篇章的方式，以助日後擬題的方向以至規劃考試課程，並從而反思增設範文對整體考生表現的影響，並檢視增設考核範文的作用，以及回饋有關結果給前線教師，藉此檢討施教的方法和效用。

初步推論，增設文言範文，使部分考生對文言篇章的理解能力提升了，故考生在課外文言篇章的表現的差異顯著擴大；至於對於理解語體文的能力未見明顯改變，宜多觀察數年的數據作分析。而增設範文，考生在甲部考題的表現差異頗大，整卷的標準差也因此增加了，有利於區分不同考生的能力。擬題時，甲部範文宜有部分考題考核高階的分析和綜合能力，除可有助區分高下外，也可以間接影響教學，甚或提升考生在不同考卷的表現。

講者簡介

葉子彬先生為香港考試及評核局中文科科目經理。
Parallel Session 4 分題研討（四）
Applications of CBT in Large Scale High-Stakes Examinations

大規模計算機化考試在全國大學英語四、六級口語考試中的應用

趙英華副處長（教育部考試中心）

摘要

全國大學英語四、六級口語考試（CET-SET）是全國大學英語四、六級考試（CET）體系中的口
語考核部分（選考），測試對象為國內全日制普通及成人高等院校本科、專科在校生，在籍研
究生，每年參加考生人數約為100萬人次。

CET-SET 全面採用計算機化考試，考生使用計算機在局域網上進行考生與模擬考官、考生與考
生之間的互動，完成考試。考試分為四級和六級兩個級別，四級考試（CET-SET4）包括短文朗
讀、簡短回答、個人陳述、雙人互動四個考核部分，六級考試（CET-SET6）分為陳述和討論、
問答兩個考核模組。CET-SET 評卷採用人工雙評，考試結束後，評卷教師在計算機上根據語言
的準確性、語言範圍、話語的長短、連貫性、靈活性、適切性等考核標準給出分數，最終以 A、
B、C、D 四個等級公佈分數。

本次交流將從 CET-SET 的考試架構設計、考試實施、考試評卷、未來發展等幾個主要方面作簡
要介紹。

講者簡介

趙英華為教育部考試中心社會考試處副處長。
Parallel Session 4 分題研討（四）
Applications of CBT in Large Scale High-Stakes Examinations

計算機輔助考試在廣東省高考英語聽說中的實踐和探索
范韶彬副院長（廣東省教育考試院）

摘要
廣東省普通高等學校招生統一考試（以下簡稱高考）是考生規模達到 70 萬人的大規模考試，為推進其中的英語科目考試的規範化、標準化，著力維護高考公平，從 2004 年起，廣東省在高考英語考試中採用了計算機技術和設備，建立了高考英語聽說計算機考試系統進行英語聽力和口語內容的考核，並開展了計算機智能評卷的試點探索。本報告主要介紹廣東省在計算機輔助高考英語聽說考試方面的基本情況。

講者簡介
范韶彬為廣東省教育考試院副院長。
Parallel Session 4 分題研討（四）
Applications of CBT in Large Scale High-Stakes Examinations

AI 在大規模高利害考試中的研究與應用

朱汝光處長（北京教育考試院）

摘要

近年來，“機器評卷”的研究不斷深入，應用實踐也日漸增多。北京教育考試院依託國內一流的人工智能研究團隊，進行了 AI 在網上閱卷質量控制、主觀題評閱、大規模英語聽說考試過程實現與結果評判等方面進行了多年的研究，並自 2018 年起，在全市中考聽說考試中進行了實質性應用，涉及考生 20 餘萬人，取得良好效果。北京教育考試院正在抓緊 AI 技術應用實證應用研究總結，不斷推進成果的應用，為新高考中的英語聽語考試提供了有益的經驗和借鑒。

講者簡介

朱汝光為北京教育考試院信息處處長。
Abstract

2020年是天津市高考综合改革的落地之年，在疫情防控常态化的背景下，考试组考工作面临著严峻的挑战。特别是口语类考试，相较于纸笔考试，其开考考场多、拾音设备重复使用的特点，都难以契合防疫的要求。天津市考试院多次就在疫情防控的情况下完成好口语类考试的方案进行了研讨。经过慎重决策、综合考虑，决定与中电科和科大讯飞两家企业进行合作，综合考试院考务、电信政务云网络以及科大讯飞“人机对话”三方的优势，以企业信息技术力量为保障，对传统考务进行有效整合，搭建了基于中电科“天翼云”和科大讯飞“人机对话”的考试平台，创新了居家线上考试和考点机房考试并行的“云考试”模式。本次“云考试”的成功是深化政企合作的一次重要尝试，也是云网融合的在线考试在大规模考试中的一次重要探索，更是发挥政企各自优势，集中社会力量办大事的制度优越性再一次得到集中体现。

Speaker Profile

范鹏为天津市教育招生考试院副研究员。
Notes 筆記
Organiser 主辦機構:
Hong Kong Examinations and Assessment Authority 香港考試及評核局

Co-organiser 協辦機構:
Education Bureau 教育局