考評局·研究論壇
HKEAA · Research Forum

45th HKEAA Anniversary

評核 • 教育 • 創未來

ASSESSMENT and EDUCATION for the future

17.11.2022

PROCEEDINGS

Table of Contents

Forum Programme (Morning Session)

Hong Kong Time (HKT)	Programme
8:45am – 9:05am	**Registration**
9:05am – 9:15am	**Opening Remarks** Mr. Samuel YUNG Wing-ki, SBS, MH, JP (Chairman, Hong Kong Examinations and Assessment Authority)
9:15am – 10:00am	**Presentation 1** ***Validity and Fairness Considerations in Innovative Assessments*** Prof. Kadriye ERCIKAN * (Vice President of Research and Measurement Sciences, Educational Testing Service, USA)
10:00am – 10:45am	**Presentation 2** ***Learning Analytics at the Intersection with Measurement for Innovative Assessments*** Prof. Kathleen SCALISE * (Professor, University of Oregon, USA)
10:45am – 11:00am	**Break**
11:00am – 11:45am	**Presentation 3** ***Technology-enhanced Large-scale Language Assessment: Opportunities and Challenges*** Prof. JIN Yan * (Professor, Shanghai Jiao Tong University, China)
11:45am – 12:30am	**Presentation 4** ***Assessment in STEM Education: Supporting and Promoting the Development of Students' Core Competence*** Prof. WANG Su * (Director of the Institute of International and Comparative Education, National Institute for Education Science, China)

Note: The research forum will be conducted in English.

* Speakers will conduct the presentation through webcast.

Forum Programme (Afternoon Session)

Hong Kong Time (HKT)	Programme
2:00pm – 2:30pm	**Presentation 5** ***Using Artificial Intelligence and Statistics to Detect Fake News and Inform Students' Media Literacy*** Prof. CHIU Ming Ming (Chair Professor, Analytics and Diversity, The Education University of Hong Kong, Hong Kong, China)
2:30pm – 3:00pm	**Presentation 6** ***Designing a Learner-first Assessment and Learning Experience to Empower Adaptive Learning*** Dr. Madeline XI Xiaoming (Director – Examinations, Assessment and Research, Hong Kong Examinations and Assessment Authority, Hong Kong, China)
3:00pm – 3:30pm	**Presentation 7** ***Using Competitions as Summative Assessment of Long-term Gifted Programmes*** Mr. TANG Wing-hong (Director, Hong Kong New Generation Cultural Association Science Innovation Centre, Hong Kong, China)
3:30pm – 3:45pm	**Break**
3:45pm – 4:15pm	**Presentation 8** ***Data speaks for Hong Kong Educators: PISA and Beyond*** Prof. HAU Kit-tai (Professor, Educational Psychology, The Chinese University of Hong Kong, Hong Kong, China)
4:15pm – 4:45pm	**Presentation 9** ***PISA Innovative Assessments of Learning Skills*** Dr. Mario PIACENTINI * (Senior Analyst, Organisation for Economic Co-operation and Development, France)

Note: The research forum will be conducted in English.

Forum Programme (Afternoon Session) (Continued)

Hong Kong Time (HKT)	Programme
4:45pm – 5:30pm	**Panel Discussion: Assessment for Future Education** **Panelists:** Prof. CHIU Ming Ming Prof. HAU Kit-tai Prof. JIN Yan* Dr. Mario PIACENTINI * Mr. TANG Wing-hong Dr. Madeline XI Xiaoming **Facilitator:** Prof. Ricardo MAK (Director – Public Examinations, Hong Kong Examinations and Assessment Authority, Hong Kong, China)
5:30pm – 5:45pm	**Concluding Remarks** Prof WEI Xiang-dong (Secretary General, Hong Kong Examinations and Assessment Authority, Hong Kong, China)

Note: The research forum will be conducted in English.

* Speakers will conduct the presentation through webcast.

Presentation Abstracts

Presentation 1

Validity and Fairness Considerations in Innovative Assessments

Prof. Kadriye ERCIKAN (Educational Testing Service)

Abstract

Innovative assessments of 21st century skills such as complex problem-solving, critical thinking and collaboration are expected to be engaging, resemble real life tasks, draw upon multidisciplinary knowledge and skills, and provide individuals with feedback on their progress towards solving problems. Several innovations are considered in large-scale assessments to meet these goals. These innovations include adaptivity based on performance on segments of the assessment, interactivity that modifies the assessment determined by student background and actions during test-taking and use of multimedia tools and digital features of assessment environments. Interactivity in assessments as well as the need for cost efficiency utilize artificial intelligence-based tools such as automated item generation and artificial intelligence-based automated scoring. While these innovations may enhance and help meet the demands for such assessments, we need to consider the new sources of threats to validity and fairness of inferences from these assessments. Implications of each of these assessment innovations and the use of technological tools on validity and fairness are discussed in this presentation. The presentation will highlight validity and fairness issues in five research and innovation projects at the Educational Testing Service: (1) a scenario based digital performance assessment originally intended for a licensure assessment; (2) an innovative digital language assessment with adaptivity and artificial intelligence based automated scoring of text and speaking; (3) use of multiple deep learning language models to automatically generate high-quality multiple choice items with keys, distractors, and metadata; (4) a personalized feedback tool for English speaking ability that can be delivered quickly and economically; and (5) caring assessments that seek to provide all students with personalized, adaptive experiences based on their motivational, social-emotional, and sociocultural characteristics and in-task behaviors. These examples are intended to provide the range of considerations in design of innovative assessments as well as types of evidence to support their use and interpretation.

About the Speaker

Kadriye Ercikan is Vice President of Research and Measurement Sciences at the Educational Testing Service (ETS) and Professor Emerita at the University of British Columbia. She is responsible for ETS's foundational and applied research and psychometric support of ETS's major testing products and contracts, and ETS's contract for the National Assessment of Educational Progress contract. Her research focuses on validity and fairness issues and sociocultural context of assessment. Her recent research includes validation of score meaning using response processes, use of response process data for comparing groups and examining fairness and validity of interpretation and use of scores from artificial intelligence based automated scoring.

Ercikan is a Fellow of the International Academy of Education, President Elect of the International Academy of Education, and President Elect of the International Test Commission. Her research has resulted in six books, four special issues of refereed journals and over 100 publications. She was awarded the American Educational Research Association (AERA) Division D Significant Contributions to Educational Measurement and Research Methodology recognition for another co-edited volume, *Generalizing from Educational Research: Beyond Qualitative and Quantitative Polarization*, and received an Early Career Award from the University of British Columbia. Ercikan is currently serving as the National Council on Measurement in Education (NCME) Book Series Editor (2021-2026).

Presentation 2

Learning Analytics at the Intersection with Measurement for Innovative Assessments

Prof. Kathleen SCALISE (University of Oregon)

Abstract

Some analytical approaches will be explored for including complex process and interaction data from technology-based products into educational assessments. An intersection is emerging between applications of learning analytics and traditional psychometric approaches to educational measurement (Papamitsiou & Economides, 2014; Scalise, Wilson, & Gochyyev, 2021). This emerging intersection may help assessments tap new technology for analyzing and integrating data reliably in innovative assessments. By establishing some common ground with good utility between fields, new approaches called "hybridized" may borrow strength across toolkits to good effect.

About the Speaker

Kathleen Scalise is a professor at the University of Oregon, with dual appointments in the Department of Education Science and in the School Psychology Program. Her main research areas are technology-enhanced assessments, data science at the intersection with measurement, dynamically delivered content in e-learning, computer adaptive and multi-stage testing, and applications to equity studies. Projects included research on 21st Century Skills with Cisco, Intel and Microsoft; Virtual Performance Assessments with Harvard University; and technology-enhanced assessments with U.S. Smarter Balanced and the National Assessment of Educational Progress (NAEP) Science.

Dr. Scalise serves internationally on Organisation for Economic Co-operation and Development (OECD)'s Program for International Student Assessment (PISA) Research Innovation Group (RIG) and previously for the International Association for the Evaluation of Educational Achievement (IEA)'s electronic version of the Trends in International Mathematics and Science Study (eTIMSS) and International Computer and Information Literacy Study (ICILS) for technology literacy. She has extensive journal publications and served on the U.S. National Research Council (NRC) report on assessment of the Next Generation Science Standards. She holds K-12 teaching credentials (California) for physical and life sciences, a B.A. in biochemistry, and the Ph.D. focusing on quantitative measurement from UC Berkeley.

Innovative contexts of interest include instructional technology especially in science and math education with simulation, gaming, e-proctoring, artificial intelligence (AI), credentialing and badges, social media, e-portfolios, and the proliferation of different types of learning evidence in big data. Her most recent book is "Why Neuroscience Matters in the Classroom," which explores the intersection of neuroscience, cognitive psychology and educational research with data (Scalise & Felde, 2017). Dr. Scalise has work in student collaboration in schools, 21st century skills such as digital literacy and problem solving, cognitive science, learning analytics, high school graduation, and policy and leadership.

Presentation 3

Technology-enhanced Large-scale Language Assessment: Opportunities and Challenges

Prof. JIN Yan (Shanghai Jiao Tong University)

Abstract

Over the past few decades, technological innovations "have taken many forms and have influenced all areas of the test development cycle, including task design, delivery, scoring, reporting and even validation" (Van Moere & Downey, 2016: 342). Innovative and useful as they are, technologies are a means, not an end. The famous quote by Cedric Price – Technology is the answer, but what was the question? – has invited educators to reflect on the nature and impact of technology on education. Donald Ely, a pioneer of educational technology, expressed his concern about overlooking the reasons for technological innovations and cautioned against the use of technology without a sound understanding of the purposes and ultimate consequences (Ely, 1995). Care therefore needs to be taken to justify technology-mediated assessment practices and balance the benefits and risks. In this presentation, I intend to showcase the technological solutions to issues involved in the development and validation of the College English Test (CET) to highlight the affordances and challenges of technology-enhanced large-scale language assessment. First, I will provide an overview of the use of technology in the CET since the late 1980s. I will then propose an evaluation framework for examining technological innovations at each stage of assessment development. Guided by the framework, I will analyze the cases of the CET Spoken English Test (CET-SET), the Internet-based CET (IB-CET), and CET automated scoring systems. Finally, I will discuss the importance of responsible use of technology that promotes "a positive transformative effect on assessment, making tests potentially more affordable, meaningful, and useful" (Schmidgall & Powers, 2017: 328).

About the Speaker

Yan Jin is a professor of linguistics and applied linguistics at Shanghai Jiao Tong University, China. She started her career as a language tester in 1991 and has been involved in the development and research of large-scale, high-stakes language testes for three decades. She is currently Chair of the National College English Testing Committee of China. She is co-editor-in-chief of the Springer open-access journal Language Testing in Asia and is also on the editorial boards of Language Testing, Language Assessment Quarterly, and a number of academic journals published in and outside China. She received the International Language Testing Association (ILTA)/Cambridge Distinguished Achievement Award in 2021.

Presentation 4

Assessment in STEM Education: Supporting and Promoting the Development of Students' Core Competence

Prof. WANG Su (National Institute for Education Science)

Abstract

With the release of a new round of national curriculum standards, school education has turned to competence -oriented education. Project-based learning and interdisciplinary learning are increasingly valued in schools. Science, technology, engineering, and mathematics (STEM) courses are also making up a larger portion of school curricula than in the past. Both scholars and schools have begun to pay attention to the assessment in STEM education, including STEM classroom assessment, STEM competencies assessment both for teachers and students. This article discusses three issues, first is about the STEM classroom assessment, and how to design an effective performance-based assessment rubric for STEM classroom. Second, STEM competence assessment for teachers, including assessment framework and assessment method. Third, STEM competencies assessment for students. The framework is developed and validity evidence is provided by empirically testing its structure.

About the Speaker

Wang Su, researcher, director of the Institute of International and Comparative Education, National Institute of Education Sciences. Her research areas focus on international comparative studies of future schools and STEM education. Since 2014, she has led the China Future School Innovation Plan, organized and developed the "Future School White Paper", "Future School 2.0 Conceptual Framework", etc., and established a network of future school research communities across the country. In 2017, the "China STEM Education 2029 Action Plan" was launched, and the "STEM Education White Paper" and "STEM competence Framework for Teacher" were organized and developed, and a network of STEM education collaborative innovation centers was established nationwide. She presided over more than ten major projects and topics of the Ministry of Science and Technology and the Ministry of Education and participated in a number of decision-making research work. She has published 30 books and more than 60 papers. She has won the second prize of the 3rd National Educational Science Outstanding Achievement Award by the Ministry of Education and the third prize of the 6th Educational Science Outstanding Achievement Award.

Presentation 5

Using Artificial Intelligence and Statistics to Detect Fake News and Inform Students' Media Literacy

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

Abstract

Fake news can kill. Many people believed COVID-19 fake news, did not protect themselves (e.g., wear mask, vaccine), got sick, and died. Thus, we test whether attributes of users or tweets can distinguish true versus fake news in 4,165 spell-checked English tweets linked to one of 20 matched COVID-19 news stories (10 true, 10 fake) via artificial intelligence, computational linguistics, and advanced statistics. Tweets with common words, negative emotional valence, higher arousal, greater dominance, first person singular pronouns, third person pronouns or by users with more followers were more likely to be true. By contrast, tweets with second person pronouns, bald starts, or hedges were more likely to be fake news. Accuracy (F1-score) was 95%. While some predictors might be universal (pronouns, politeness, followers), others might be topic-specific (common words, emotions, hedges). We will model diffusion scope, speed and shape of fake news to build a student dashboard.

About the Speaker

Ming Ming CHIU is *Chair Professor of Analytics and Diversity* and Analytics\Assessment Research Center *Director*, The Education University of Hong Kong. A graduate of Columbia BS (computer science), Harvard EdM (interactive technology) and UC-Berkeley PhD (education), he was *Senior 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 (top 50 learning science idea –*International Society of the Learning Sciences*), (b) *multilevel diffusion analysis* to detect corruption in the music industry, (c) artificial intelligence *Statistician*, and (d) online detection of sexual predators. His 77 grants (US$16 million) yielded 267 publications (178 journal articles; 10,000+ citations; #8 in Education in China, 2020), 13 keynote speeches, 5 television broadcasts, 17 radio broadcasts, and 170 news articles in 21 countries. He creates automatic statistical analyses for Big Data.

Presentation 6

Designing a Learner-first Assessment and Learning Experience to Empower Adaptive Learning

Dr. Madeline XI Xiaoming (Hong Kong Examinations and Assessment Authority)

Abstract

Individualization, customization and adaptivity have become the catchwords in education. However, learner-first adaptive learning solutions, where a learner's needs and wants are prioritized every step of the way when he/she interacts with the assessments and learning content, are rare. This is because developing such solutions requires interdisciplinary talents in assessment, learning, cognitive and non-cognitive science, artificial intelligence (AI), and many more, which, in reality, is a luxury for most development teams.

How do we ensure a learner-first assessment and learning experience? In designing various types of assessments in adaptive learning, we want the assessments to be efficient yet precise, unobtrusive, provide actionable information, and support a positive assessment taking experience. A learning experience optimized for an individual learner must meet his/her unique learning needs, and be tailored to his/her level, dynamic knowledge and skill profiles, cognitive and learning styles, and constantly changing affective states to facilitate the most speedy and effective learning.

In this talk, I will discuss the science and technologies behind an adaptive learning system and decompose its architecture, focusing on the chain of inferences supporting its overall efficacy, including user property representation, user property estimation, content representation, user interaction representation, and user interaction impact. I will provide an overview of different types of assessment used in adaptive learning and an analysis of the assessment approach, priorities, and design considerations of each to optimize its use in adaptive learning. I will then propose a framework for evaluating different aspects of an adaptive learning system. I will conclude with thoughts on high-priority research and development to provide truly learner-first systems to fully empower our learners.

About the Speaker

Dr Xi Xiaoming Madeline is the Director – Examinations, Assessment and Research. She leads the Assessment Technology and Research Division, the International and Professional Examinations Division and the Education Assessment Services Division, and is in charge of formulating strategies and policies governing research initiatives, promoting assessment for learning, conducting analytics in support of the standards of the Hong Kong Diploma of Secondary Education Examination (HKDSE), operating and improving the Territory-wide System Assessment (TSA) project and exploring research/business partnership and collaboration opportunities with external parties.

Dr Xi is an internationally recognised expert in assessment, the interface between learning and assessment, and educational technologies including AI technologies with expertise in educational product research and development (R&D), strategy, implementation and growth. She has substantial research experience and published widely on testing, assessment, learning, and educational AI technologies.

Prior to joining the HKEAA in April 2022, she had a long tenure with the Educational Testing Service (ETS) where she was the Senior Research Director of the Centre for Language Learning and Assessment and subsequently the Executive Director, New Product Development when she left in 2020. She had then taken up the role of Chief of Product, Assessment and Learning in VIPKID. Dr Xi obtained her Doctor of Philosophy in Educational Measurement and Statistics and Language Assessment from the University of California, Los Angeles and Master of Business Administration from the University of North Carolina at Chapel Hill.

Presentation 7

Using Competitions as Summative Assessment of Long-term Gifted Programmes

Mr. TANG Wing-hong
(Hong Kong New Generation Cultural Association Science Innovation Centre)

Abstract

Assessment to gifted students is always difficult. How to define a student is Scientifically, Technologically or Innovatively gifted, is also a great problem. This is to share the work of Hong Kong New Generation Cultural Association (HKNGCA) Science Innovation Centre, one of the two Education Bureau (EDB) committed Gifted Education Programs providers in the past 16 years. General statistics of the work done with selected cases will be shared to show the achievements of our secondary students in the International Science and Innovation competitions.

About the Speaker

Mr. TANG Wing-hong is the Director of the Science Innovation Centre of the Hong Kong New Generation Cultural Association. He holds the following education qualifications: MA(LIBS) HKUST, BEd(Hon)(CWLD), Teach Cert(D&T)HKTTC. He was the former Vice Principal (Curriculum & Students' Development) of CCC Rotary Secondary School, a member of the CDCC on Technology Education (1999-2006), the Chairman of the HKEAA Subject Committee on Science & Technology (HKCEE) (2000-2003), and the Chairman, HKEAA/CDCC Joint Committee on Design and Applied Technology (HKDSE) (2003-2013), Teaching Consultant of Education University of Hong Kong, and the recipient of the 2008-2009 Chief Executive's Award for Teaching Excellence (CEATE Technology Education). Among his social services is his chairmanship of the Board of Directors of the Tsuen Wan Government Secondary Technical School Alumni Association Charitable Foundation Limited.

Presentation 8

Data speaks for Hong Kong Educators: PISA and Beyond

Prof. HAU Kit-tai (The Chinese University of Hong Kong)

Abstract

Obviously, gender differences cannot be examined with girls only. Similarly, despite various challenges, large scale comparison studies across many economies become very important in order to understand the characteristics of students in a certain city or country. Results from various large-scale educational surveys (PISA by OECD, TIMSS by IEA, TALIS by OECD, TSA by HK EDB) will be used to show the strengths and weaknesses of Hong Kong students with reference to other high academic performing systems. Various myths such as the importance of e-learning and project learning, as well Hong Kong or Asian students' low self-concept, meaning of life and high anxiety, will be discussed.

About the Speaker

Kit-Tai Hau was formerly Choh-Ming Li Professor of Educational Psychology and vice-President, The Chinese University of Hong Kong. He is Fellow of the American Educational Research Association and the International Applied Psychology Association; member of the Organisation for Economic Co-operation and Development (OECD) Programme for International Student Assessment (PISA) Technical Advisory Group, Strategic Development Group and Questionnaire Expert Group, and member of the Trends in International Mathematics and Science Study (TIMSS) Questionnaire Item Review Committee.

Presentation 9

PISA Innovative Assessments of Learning Skills

Dr. Mario PIACENTINI (Organisation for Economic Co-operation and Development)

Abstract

Traditional assessments have served well the purpose of measuring to what extent students have acquired discrete pieces of knowledge. However, in rapidly changing societies, students' future will depend not only on the knowledge they have acquired during schooling but also on their ability to learn new things. It is thus important to develop new assessments that focus on learning skills. This speech will present the PISA 2025 assessment of Learning in the Digital World, that has been designed to produce internationally comparable evidence on students' capacity to construct their own knowledge and understanding with digital learning environments. All units in this new test are designed as online tutoring experiences of about 30 minutes, where students learn with a virtual tutor how to use digital tools to solve problems or explore systems. Metrics on learning skills are developed by comparing students' responses to a pre-test to their performance on the learning tasks, and by analyzing how students use the learning resources. The presentation will also introduce some innovative assessments with learning affordances the OECD has been developing for formative use in the classroom.

About the Speaker

Mario Piacentini is a senior analyst in the Programme for International Student Assessment (PISA) at the OECD. An expert in measurement, Mario leads the work on PISA innovative assessments and the PISA Research and Development (R&D) Programme. He works with international experts to design assessments of 21st century competences. His projects aim to expand the metrics we use to define successful education systems.

He is the lead author of the Global Competence (PISA 2018) and Creative Thinking (PISA 2022) assessment frameworks. He is now leading the development of the PISA 2025 assessment of Learning in the Digital World. He also coordinates a research project on technology-enhanced, formative assessments.

Before joining PISA, he worked for the Public Governance and the Statistics Directorates of the OECD, the University of Geneva, the World Bank and the Swiss Cooperation. He authored several peer-reviewed articles and reports. Mario holds a PhD in economics from the University of Geneva.

Notes

Notes

Proceedings:

Questions:

Evaluation Form:

Website:
http://www.hkeaa.edu.hk

Email:
rf2022@hkeaa.edu.hk

Organiser:

Hong Kong Examinations and Assessment Authority