Opportunity versus Challenge:

Exploring Usage of Log-File and Process Data in International Large Scale Assessments

Conference and Workshop Programme

16th – 17th May 2019
Hotel Riu Plaza the Gresham,
Dublin 1, Ireland

ETS®

Educational Research Centre
Foras Taighde ar Oideachas

Conference and workshop overview

The Educational Research Centre (Dublin, Ireland) and Educational Testing Service (Princeton, NJ, USA) are collaborating to support the growing interest in log-file and process data from international large-scale assessments. The conference event **Opportunity versus Challenge: Exploring Usage of Log-File and Process Data in International Large Scale Assessments** takes place on the 16th and 17th May, 2019 at Hotel Riu Plaza the Gresham in Dublin, Ireland.

The use of log-file and process data is a recent and evolving area of educational research. By nature, this field is complex due both to the structure of log-file data and also to the complexities in analysing and interpreting the data. Yet, despite these challenges, significant work has been achieved, and exciting and innovative exploratory work, both data-driven and theory- or model-driven, are underway. Further, the availability of log-file data raises important ethical questions: for example, should this data be used to assign proficiency scores to test-takers, or should it be limited to explaining variations in proficiency, or should its remit be limited to data and measurement quality?

The conference aims to:

- Showcase recent and current work in this area that takes a variety of methodological and theoretical approaches

- Promote a shared understanding of log-file and process data in PISA and other large-scale assessments

- Consider ethical issues relating to this kind of data

- Establish priorities to guide future directions for research using log-file data.

The workshop aims to:

- Cover the extraction process and data structure

- Combine demo/hands-on tasks for those starting to work with process data

- Outline some key manipulations.

The workshop will use a synthetic database of Field Trial data from three countries/economies (Germany, Ireland and Singapore) that participated in PISA 2015.

Conference website:
www.erc.ie/processdataconference2019

Conference hashtag:
#ProcessDataConf2019

To contact the conference and workshop team, please email:
processdataconference2019@erc.ie

Conference outcomes:
The conference presentations will be published as a special issue of a journal.

Day 1 Conference 9.00 – 17.15, Thursday 16th May

Time	Session	Speaker
8.30 – 9.00	**Registration, tea and coffee**	
9.00 – 9.30	**Welcome** ■ Context for and aims of the conference	Irwin Kirsch, *ETS*
9.30 – 10.30	**Keynote address** ■ Process data, the new frontier for assessment development: Rich new soil or a Quixotic quest?	Stephen Provasnik, *NCES*
10.30 – 11.00	**Tea/Coffee break**	
11.00 – 12.30	**Session 1 – Using process data to go beyond cognitive test scores: Examples and challenges** ■ Using timing information associated with response data in large-scale assessments ■ Effort and achievement of 15-year-olds in PISA 2015 across EU member states	Kentaro Yamamoto, *ETS* Loris Vergolini, *IRVAPP*
12.30 – 13.30	**Lunch (served in Gallery Restaurant)**	
13.30 – 15.00	**Session 2 –Theory-driven applications of process data to identify behavioural response patterns** ■ Log file analyses, transversal skills, and learning analytics: A good match, but not a perfect one ■ Exploring sequence-based approaches using process data in large-scale assessments	Samuel Greiff, *University of Luxembourg* Qiwei He, *ETS*
15.00 – 15.30	**Tea/Coffee break**	
15.30 – 16.15	**Session 3 – Ethical issues related to process data analysis** ■ Ethical considerations involving data analytics in educational assessment: A systematic literature review	Damian Murchan, *Trinity College Dublin*
16.15 – 17.00	**Session 4 – Issues of validation** ■ Some considerations in validating the interpretation of process indicators	Frank Goldhammer, *DIPF*
17.00 – 17.15	Closing remarks	Jude Cosgrove, *ERC*
17.15 – 18.30	**Drinks reception – Gallery Restaurant, the Gresham Hotel** (Conference participants organise their own evening meal – there are many nearby places to eat)	

Day 2 Conference 9.00-13.00, Friday 17th May

Time	Session	Speaker
8.30 – 9.00	**Registration for the data workshop**	
9.00 – 10.30	**Session 5 – Using process data for validation: domain constructs, cognitive processes and assessment frameworks** ■ Using national assessment process data to identify construct relevant variables and infer cognitive processes ■ Selection of irrelevant navigation paths during digital reading: Underlying cognitive processes	Luis Saldiva, *ETS* Ladislao Salmerón, *University of Valencia*
10.30 – 11.00	**Tea/coffee break**	
11.00 – 11.45	**Session 6 – Data-mining applications of process data to identify behavioural response patterns** ■ Assessment in the age of data science: The case of interactive items tested in France	Reinaldo Dos Santos, *Ministère de l'éducation nationale, de l'enseignement supérieur et de la recherche*
11.45 – 12.45	**Session 7 - Interactive session to include a Q&A / conference feedback for next steps** Interactive, moderated discussion in three streams: ■ Conceptual ■ Technical/methodological ■ Ethical/data protection	Facilitated and moderated by *ERC*
12.45 – 13.00	Conference closing session	Stephen Provasnik, *NCES*
13.00 – 14.00	**Lunch (served in Gallery Restaurant)**	

Day 2 Workshop 14.00-18.00, Friday 17th May

Time	Session	Workshop Facilitators
14.00 – 14.15	Welcome, overview, and objectives	
14.15 – 15.30	**Workshop session 1** ■ Overview of the extraction and pre-processing data from the PISA 2015 .xml files, structure of files, codebook	Eugenio Gonzalez & Qiwei He, *ETS*
15.30 – 15.45	**Tea/coffee break**	
15.45 – 17.30	**Workshop session 2** ■ Demo and hands-on of general manipulations of the data driven by research questions	
17.30 – 18.00	Workshop closing session: Q&A	

Abstracts of Conference Presentations

Keynote

Process data, the new frontier for assessment development: Rich new soil or a Quixotic quest?

Stephen Provasnik, NCES

This address provides a context and common set of definitions of key concepts for the subsequent discussions and presentations on log files and process data. It also provides an introductory overview of current uses of process data, and considers what we might reasonably use process data for. The presentation concludes with some thoughts on where we might go (and, perhaps go wrong!) by venturing further into this new 'process' land.

Session 1 – Using process data to go beyond cognitive test scores: Examples and challenges

Presentation 1: Using timing information associated with response data in large-scale assessments

Kentaro Yamamoto, ETS

Timing data can be used to increase the reliability and validity of measurement that supports the measurement framework beyond which traditional cognitive data can provide. Response timing often retains artifacts of data fabrication by various players including the respondent, administrator, data managers and so on. It can also indicate the amount and type of errors included in the response data, such as the interpretation of omitted responses in regard to how much it reflects the proficiency of the respondent. Further, it can support the interpretation of measurement constructs. This presentation will explore these issues with examples. Response timing and proficiency have complex relationships and these relationships are strongly contextual and construct specific, thus psychometric modeling of timing and proficiency and its generalisations should be made with caution.

Presentation 2: Effort and achievement of 15-year-olds in PISA 2015 across EU member states

Loris Vergolini, IRVAPP (presenting), with
Davide Azzolini (FBK-IRVAPP), Nicola Bazoli (FBK-IRVAPP), Christian Monseur (University of Liége), Elodie Pools (University of Liége), and Antonio Schizzerotto (University of Trento and FBK-IRVAPP)

The aim of this paper consists in the analysis of the relationship between effort and performance using the PISA 2015 data. To measure effort, we exploit the fact that the tests in PISA 2015 have been administered via computer. This procedure generates the so-called log-files that are the traces left by students in answering the test that can be used to measure non-cognitive competencies. Effort is measured as the difference in the response time between difficult and easy items in a given cluster of the booklet, assuming that a student put effort in completing the test if he/she takes enough time to solve more complex items. For what concerns the relationship between effort and performance, we will focus on the possible influence of effort on the three domains separately (i.e., science, maths and reading). We will apply OLS regression to estimate this influence at micro-level controlling for a wide range of individual (i.e., social origins, gender, migration background, etc.) and school level (extra-curricular activities, school climate, etc.) characteristics. The analytical strategy is built on two steps. First, we estimate a pooled model with all the countries to understand if a relationship between effort and achievement exists at EU level. Second, we estimate separate models for each country to assess if and how this relationship varies between countries.

Session 2 –Theory-driven applications of process data to identify behavioural response patterns

Presentation 1: Log-file analyses, transversal skills, and learning analytics: A good match, but not a perfect one

Samuel Greiff, University of Luxembourg

The 21st century has brought along a number of societal and technological changes that have had a profound impact on globally operating educational large-scale assessments. Three of those changes are addressed in this presentation: (1) the move towards computer-based assessments even on a large scale that has led to the availability of large amounts of computer-generated log files on single student actions; (2) the broader understanding of the term educational attainment that has led to the inclusion of transversal skills (sometimes called innovative domains) in large-scale assessments such as interactive problem solving in PISA 2012 or collaborative problem solving in PISA 2015; (3) a general move from assessment of learning (summative assessment) to assessment for learning (formative) leading to methods that support and facilitate the learning experience even in an assessment situation. This presentation uses PISA to highlight how these three developments are intertwined and what implications they have for the future of large-scale assessments. It will draw on recent findings in the field of interactive (PISA 2012) and collaborative (PISA 2015) problem solving to show how log-file analyses and transversal skills go hand-in-hand and how they can lead towards the integration of learning analytics and components of formative assessment in the understanding of student performance and in the reporting of studies such as PISA. In fact, the recent developments and the comprehensive shifts conducted in PISA and beyond might ultimately lead to a revolution in the understanding of processes and behavioural patterns and in the way educational large-scale assessments are internationally utilised.

Presentation 2: Exploring sequence-based approaches using process data in large-scale assessments

Qiwei He, ETS

Identifying sequential patterns in process data can be useful for discovering, understanding, and, ultimately, scaffolding test takers problem solving behaviours. Sequence-based approaches developed in natural language processing and biostatistics were motivated to be applied in process data that have similar structures as language, though more complex. This presentation will mainly introduce two sequence-based approaches using process data collected in problem solving items in Programme for International Student Assessment (PISA) and Programme for the International Assessment of Adult Competencies (PIAAC). The first approach features in disassembling long sequences into mini-sequences by n-grams. These n-gram-based features could be further selected as the robust classifier to distinguish different subgroups. The second approach takes the sequence as a whole. It focuses on identifying the longest common subsequence between the individual action sequence and the pre-defined ones by calculating the distance in between. Both of these approaches hold promise in process data analysis from different perspectives. The first approach is more favourable to the item-level analysis given the action sequences are often dependent on the particular task being analysed, while the second approach provides the possibility to generalize factors that are associated with test takers' problem-solving behaviours across multiple items.

Session 3 – Ethical issues related to process data analysis

Ethical considerations involving data analytics in educational assessment:
A systematic literature review

Damian Murchan, Trinity College Dublin (presenting), with
Fazilat Siddiq, Department of Education and Quality in Learning, University of South-Eastern Norway

Computer-based assessments (CBAs) have potential to provide valuable information for stakeholders, especially in the context of the digital agenda in education. Benefits of CBAs include flexibility, efficiency, consistency in scoring and increased student engagement. Another advantage is potential for employment of data mining and data analytics techniques to model examinees' engagement with items, including a variety of simple and more complex tasks. Approaches to data analytics have captured the imagination of educational researchers, even if on-the-ground illustrations lag somewhat behind. Optimism about the power of process data to help explain and formatively guide student learning is tempered somewhat by technical challenges. This study explores another challenge, namely ethical and privacy issues surrounding the use of log-file data for purposes that may not have been adequately explained to the students. Set against a backdrop of accelerating advances in computational processing, public concern about the privacy of online data and recent EU data regulations, this systematic review of literature audits practice in log-file analysis, provides up-to-date knowledge, and uncovers research gaps. Highlighting the extent to which issues of privacy, consent, individuals' rights and ethics have been reflected in research using process data drawn from student assessments, the findings provide useful evidence to assessment professionals and help inform future research and practice.

Session 4 – Issues of validation

Some considerations in validating the interpretation of process indicators

Frank Goldhammer (presenting), DIPF, with
Caroline Hahnel (DIPF), Ulf Kroehne (DIPF), and Fabian Zehner (DIPF)

International large-scale assessments such as PISA or PIAAC have started to provide public use files for log data; that is, events, related attributes, and timestamps stored when test takers interact with the assessment system. Log data and process indicators calculated from it can be used for many purposes. In educational measurement, for instance, they may be used for tapping new process-related constructs, for supporting quality control, or for optimising the test design. Furthermore, they may help to validate test score interpretations and to address substantive research questions. However, the intended uses and interpretations of process indicators require validation; that is, a theoretical and/or empirical justification that inferences about ascribed (latent) states of information processing are valid. This presentation will shed light on how to assure a valid interpretation by an evidence-centered design of the tasks and the assessment system, and on how to empirically challenge the intended interpretation of process indicators by developing and implementing correlational and/or experimental validation strategies. This could include, for instance, adapting the classical approaches of nomothetic span and construct representation or methodological triangulation under consideration of quantitative and qualitative data. We will show that contextualising information plays an important role when defining or using process indicators in order to reduce interpretative ambiguities.

Session 5 – Using process data for validation: domain constructs, cognitive processes and assessment frameworks

Presentation 1: Using national assessment process data to identify construct relevant variables and infer cognitive processes

Luis Saldiva, ETS

As students interact with an item or an interactive component within an item, they leave a trail of observable events, which can potentially be used to make inferences about their cognitive processes. While events are a salient and intuitive way to think about observable processes, it is also helpful to consider a complementary representation, namely the problem states, or states for short. These are construct-relevant (and possibly irrelevant) variables in the problem that students potentially have to keep track of. Examples of events and states generated by the students' work on items will be presented and construct relevant inferences based on these events will be discussed in this presentation.

Presentation 2: Selection of irrelevant navigation paths during digital reading: Underlying cognitive processes

Ladislao Salmerón, University of Valencia

A major component of digital reading is the ability to move within different complex structures of Internet pages. Accordingly, savvy digital readers are expected to navigate across pages that contain information relevant for the question to be answered, or the task to be solved, while staying away from the potentially endless number of pages with irrelevant information. Indeed, students with good reading comprehension skills visit more relevant paths, and in turn this improves their performance. The interpretation of visits to irrelevant pages has received less attention in the literature, as it consider to represent students' 'lostness'. But visits to irrelevant pages may also represent an attempt to fully explore the digital text to capture a complete view of the theme. This aspect should be more obvious in digital texts that are highly topically interconnected, such as Wikipedia texts, than in scenarios with highly differentiated sections, such as web portals. We tested these predictions in two datasets of 7th-10th grade students that completed a digital reading test inspired by PISA. Preliminary results provide evidence for our hypotheses. Results emphasize the need to consider irrelevant visits as an indicator of different underlying cognitive processes, depending on the interconnectedness of the digital reading scenario.

Session 6 – Data-mining applications of process data to identify behavioural response patterns

Assessment in the age of data science: The case of interactive items tested in France

Reinaldo Dos Santos, Ministère de l'éducation nationale, de l'enseignement supérieur et de la recherche

The culture of assessment is different from one country to another. The French educational system is still influenced by classroom evaluation. Teachers are responsible for their own ways to proceed to summative or formative assessments. Standardised testing often suffered from its lack of feedback in the learning process. To improve this feedback, we need to be able to observe the actions realised by the students during the test. Computer-based assessments give us the ability to go beyond the students' response. We can also collect the actions realised by the students to find the answer to the item. We can then analyse the collected data to figure out the strategies developed by the student to solve the item. By characterising the students' errors, the assessment can allow the teachers to better understand the cause of the error, and then to adapt the learning process. Accordingly, an analysis on three interactive items of mathematics, included in a national assessment for students in 9th grade, is presented in this paper.

Speaker/Facilitator Biographies

Jude **COSGROVE** is CEO of the ERC since September 2018. She has worked on national and international assessments including PISA, PIAAC and ICCS in various roles for over 20 years, and currently represents Ireland on the Governing Board for PIAAC. Her PhD examined the validity of national policy claims made on the basis of international educational assessment results. She has a keen interest in digital technologies in education, and is currently overseeing a national evaluation of the Digital Learning Framework, part of Ireland's *Digital Strategy for Schools 2015-2020*.

Reinaldo **DOS SANTOS** is a data analyst for the French Ministry of Education. He holds a Master degree in Physics (Pierre and Marie Curie University, 2006) and a Bachelor degree in Social Sciences (Paris Nanterre University, 2007). He is involved in the analysis of standardised tests performed in French institutions, particularly in process management and methodology. Using his skills in psychometrics and automatic learning, he participates in innovative research made possible by computerised assessments.

Frank **GOLDHAMMER** is head of the Centre for Technology Based Assessment (TBA) at the DIPF (Leibniz Institute for Research and Information in Education), and Full Professor of Educational and Psychological Assessment (Technology-Based Assessment and Instruction) at Goethe University Frankfurt a.M. and the Centre for International Student Assessment (ZIB). He holds a PhD in psychology from the Goethe University Frankfurt a.M. Germany. His extensive research includes technology-based assessment (e.g., validity issues), the use of process data for assessment purposes, statistical modelling and experimental control of response times, digital skills and motivational context variables. Dr Goldhammer has been continuously involved in educational large-scale assessments both at the national and international level (e.g., PISA 2009-2018, PIAAC 2012, ICILS, TIMSS).

Eugenio **GONZALEZ** has a vast and varied experience in the field of large-scale assessments in education. He is currently Principal Research Project Manager at ETS, and the Director of the IEA-ETS Research Institute (IERI), a collaborative effort between the IEA and ETS that focuses on improving the science of large-scale assessments. His previous positions include Head of the IEA's Research & Analysis Unit, and Director of Quality Control and Field Operations at the US National Assessment of Educational Progress (NAEP). While working on his doctorate at Boston College, he oversaw the development and implementation of TIMSS and PIRLS.

Dr Gonzalez, along with Dr He, are facilitators of the data workshop at this event.

Samuel **GREIFF** is head of research group, principal investigator, and Full Professor of Educational Assessment and Psychology at University of Luxembourg. Prof Greiff has been awarded several national and international research funds (overall funding approx. €9.3M), was fellow in the Luxembourg research programme of excellency, and has published articles in national and international scientific journals and books (>100 contributions in peer-reviewed journals). He has an extensive record of conference contributions and invited talks (>200 talks) and serves as editor for several journals including the *European Journal of Psychological Assessment*, as associate editor for *Intelligence*, and as guest editor for *Journal of Educational Psychology, Computers in Human Behavior*, and *Journal of Business & Psychology*. He currently serves on five editorial boards. He has been involved in PISA since the 2012 cycle as external advisor to the PISA 2012 and 2015 Expert and Subject Matter Expert Groups. He serves also as chair of the problem solving expert group for the 2nd cycle of PIAAC. Prof Greiff has been working for several years on the assessment of complex and collaborative problem solving and their role in the classroom, at work, and in private life. He has a strong interest in the use of log-file data and learning analytics for summative and formative assessment purposes.

Qiwei (Britt) **HE** is a Research Scientist in the Center for Next Generation Psychometrics and Data Science at Educational Testing Service (ETS). Her research focus is situated in the field of educational and psychological measurement, with specific attention to methodology advancement in large scale assessments (e.g., PISA, PIAAC) and complex data source (e.g., process data, textual data). The innovative nature of her research has received high recognition, including the 2019 Jason Millman Promising Measurement Scholar Award by the US National Council on Measurement in Education (NCME); 2017 Alicia Cascallar NCME Award for an Outstanding Paper by an Early Career Scholar; and acceptance into the OECD Thomas J. Alexander Fellowship (2017-2018). She is also leading an NSF-funded project as one of the principal investigators to develop latent and graphical models for complex dependent data in education. Dr He received her PhD in Psychometrics and Data Analysis at the University of Twente, Netherlands in 2013. Her doctoral research focused on combining techniques in text mining and item response modeling for handling both unstructured and structured data automatically in psychiatric and psychological assessments.

Dr He is both a presenter and, along with Dr Gonzalez, is one of the two workshop facilitators at this event.

Irwin **KIRSCH** is the Ralph Tyler Chair in Large Scale Assessment and Director of the Center for Global Assessment at ETS in Princeton, NJ. In his role as Director of the Center he oversees several teams of research scientists and project managers who are responsible for the development, management and implementation of various large scale international assessments as well as the creation of policy reports focusing on human capital and education. Over the course of his career, Dr Kirsch has worked in close collaboration with a number of state, national and international organisations including the World Bank, UNESCO, IEA and OECD, and he currently serves as the International Survey Director for both the PISA and PIAAC programmes of work.

Damian **MURCHAN** is Head of the School of Education in Trinity College Dublin. His research interests include assessment methods and reform in education, e-learning and assessment, incorporation of key skills into the curriculum, and teacher professional development. He is co-author of a recent SAGE book, *Understanding and Applying Assessment in Education*. He has held a number of advisory roles in relation to the development of assessment policy and practice in Ireland and internationally.

Stephen **PROVASNIK** leads the International Activities Program at the US Department of Education's National Center for Education Statistics (NCES). He has served as the US National Research Coordinator for the Trends in International Mathematics and Science Study (TIMSS) and Progress in Reading Literacy Study (PIRLS). He was a national programme manager for cycle one of the Programme for the International Assessment of Adult Competencies (PIAAC). He has been the US representative to the International Association for the Evaluation of Educational Achievement (IEA) governing board since 2017 and the PIAAC Board of Participating Countries since 2016. Before joining the International Activities Program he served as NCES deputy editor for *The Condition of Education,* which provides an annual statistical profile of the state of education in the United States. In his career, he has worked variously as a public high school teacher in Japan, reporter for Gannett and the *Yomiuri Shimbun,* historian of education, assessment developer for North Carolina's and Ohio's state social studies assessments, and researcher at the American Institutes for Research. He earned his doctorate in History from the University of Chicago.

Luis **SALDIVA** is the Executive Director for the National Assessment of Educational Progress (NAEP) in the Assessment Division at ETS. Before joining the ETS-NAEP team, Dr Saldivia led the development of different assessments including the Advanced Placement (AP) Statistics Test, the Graduate Record Examination (GRE) Mathematics Subject Test, the College-Level Examination Program (CLEP) College Algebra Test, and the Examen de Admisión a Estudios de Postgrado (EXADEP). He also led the ETS team that helped translating the Common Core State Standards-Mathematics (CCSS-M) for assessment purposes for the Partnership for Assessment of Readiness for College and Careers (PARCC). Dr Saldivia co-led the ETS research initiative *Test Development Quality and Efficiency* and has been involved in several research projects including projects related to automatic item generation, predicting the effect of internal factors on the discrimination power of items, estimating item difficulty with comparative judgment, and investigating the effect of item design decisions on the cognitive, affective, behavioural, and social processes. Dr Saldiva received his PhD in Mathematics from Michigan State University in 2003.

Ladislao **SALMERÓN** is associate professor and director of the Reading Institute of the University of Valencia, Spain. He obtained his PhD in 2006 on experimental psychology at the University of Granada, Spain, and was a Fulbright visiting scholar at the Institute of Cognitive Science of the University of Colorado, USA, during 2007. His research focuses on the cognitive mechanisms of digital reading. Within this field, he has studied the development of digital reading competence across adolescence, digital reading skills of students with intellectual disability, and the use of eye-movement videos to support the acquisition of digital reading abilities.

Loris **VERGOLINI** is currently a research fellow at FBK-IRVAPP, a research institute of the Bruno Kessler Foundation (Italy). His research work lies at the intersection between policy evaluation and social inequalities. More precisely, his specific areas of expertise are the evaluation of educational policies, the consequences of anti-poverty measures, and the analysis of the inequalities of educational outcomes. Loris and colleagues received funding from the European Commission to examine student engagement and achievement using PISA 2015 log-file data. Published in 2019, the policy relevance of the study lies in establishing the role played by the most prominent factors of education inequality (i.e. social background, gender, and migration background), and in identifying the main school-level and education system factors that foster non-cognitive skillsets in the young population.

Kentaro **YAMAMOTO** is a principal research scientist for Global Assessment at Educational Testing Service (ETS). Dr Yamamoto has been a technical advisor for the OECD and US Department of Education. Over the past 25 years, he has designed, or contributed in designing, numerous national and international large scale surveys of various subject domains for adults as well as for school-going populations such as NAEP, TIMSS, PISA, IALS, ALL, and PIAAC. He has also designed several individual tests in reading and literacy. Dr Yamamoto developed a psychometric model called Hybrid, a mixture model of continuous and discrete measurement models for diagnostic testing and IRT scaling, and this has been used for all literacy surveys at ETS. He also designed the online testlet adaptive testing for Prose, Document and Quantitative adult literacy skills associated with PIAAC. He has written numerous reports, research papers, and technical reports, contributed chapters in multiple books and has given numerous presentations at the national and international conferences.

Thanks

We would like to thank the participants at a symposium event held in ETS, Washington DC, in December, 2018. A special thanks to Qiwei He, ETS, for leading on the organisation of this symposium. The discussions in Washington DC proved invaluable in organising this conference. Participants were:

Ruhan Circi, AIR; Jude Cosgrove, ERC; Sylvia Denner, ERC; Matthias von Davier, NBME; Kadriye Ercikan, ETS; Frank Goldhammer, DIPF; Pauline Givord, OECD; Qiwei He, ETS; François Keslair, OECD; Irwin Kirsch, ETS; Jingchen Liu, Colombia University; Dan McGrath, NCES; Caroline McKeown, ERC; Zbigniew Marciniak, Warsaw University; Piotr Mitros, ETS; Jonathan Osborne, Stanford University; Marco Paccagnella, OECD; Stephen Provasnik, NCES; Jean-François Rouet, CNRS, University of Poitiers; Luis Saldiva, ETS; Emmanuel Sikali, NCES; Claudia Tamassia, ETS; William Thorn, OECD; and Kentaro Yamamoto, ETS.

This conference is a collaborative effort between ETS and the ERC.
We would like to extend a special thank you to:

- Caroline McKeown, ERC, who led the organisation of the conference

- Qiwei He, ETS and Eugenio Gonzalez, ETS, who led the organisation of the workshop

- Peter Archer, ERC (CEO up to September 2018), Irwin Kirsch, ETS, and Gerry Shiel, ERC, for guiding symposium and conference planning

- Paula Chute, ERC (EO up to January 2019), Anne Comey, ERC, Brenda Donohue, ERC, Alice Duggan, ERC, Adrian O'Flaherty, ERC, Imelda Pluck, ERC, and Judith Shahbazian, ETS, for administrative and technical support

- Mike Wagner, ETS, for data extraction and structuring for the workshop datafiles

- The national centres in Germany, Ireland and Singapore, who granted access to the .xml log files from the PISA 2015 field trial. These files were used to construct the synthetic database of released science items for the workshop.