Invited Symposium:
What are PIAAC process data used for?

Chairs
Frank Goldhammer, Ulf Kröhne & Carolin Hahnel (DIPF | Leibniz Institute for Research and Information in Education, Germany)

Abstract
The first round of the PIAAC study 2011–2012 was innovative in many ways. It was not only the first international large-scale assessment to be mainly computer-based, but it was also the first large-scale assessment that systematically collected log data, that is, events, event-related attributes, and timestamps reflecting the test taker's interactions with the PIAAC assessment system. Log data found their way into the public use file in the form of generic process indicators such as time on task or number of interactions by item. Moreover, the majority of countries participating in PIAAC round one provided their log data that was recorded during the computer-based assessment to make it publicly available to the research community. Together with the result data and the data from the background questionnaire, the PIAAC log data already has inspired intensive research activities (for an overview see Goldhammer et al., 2020). In this invited symposium, we have brought together four recent contributions dealing with PIAAC process data from different – both substantive and methodological – perspectives. The first presentation by Pokropek and colleagues addresses the question of how process data can be used to explore gender differences in literacy as assessed in PIAAC and PISA. The following contribution by Hahnel and colleagues investigates strategies of information processing in simulated web search environments included in the PIAAC assessment of problem solving in technology-rich environments. The third presentation by He and colleagues focuses on evaluating consistency of adult behavioral patterns across multiple problem solving tasks using PIAAC process data from the US. The final presentation by Maddox is about the PIAAC testing situation at home and considers process data derived from video data, log data etc. to capture sources of variation in the test administration.

Presentations
1. Title: Utilizing process data to examine gender differences in literacy in PISA and PIAAC.
   Authors: Artur Pokropek (Institute of Philosophy and Sociology of the Polish Academy of Sciences, Poland), Francesca Borgonovi (University College London, UK, and OECD, France), & Lale Khorramdel (Boston College, USA)
2. Title: Identifying strategies of information processing in PIAAC web search environments.
   Authors: Carolin Hahnel, Frank Goldhammer, & Ulf Kroehne (DIPF | Leibniz Institute for Research and Information in Education, Germany)
3. Title: Evaluating consistency of adult problem-solving behaviors across multiple tasks using PIAAC process data.
   Authors: Qiwei He (Educational Testing Service, USA), Dandan Liao (Cambium Assessment, Inc, USA), Hok Kan Ling (Queen's University, Canada), & Hong Jiao (University of Maryland, USA)
4. Title: Testing situations at home.
   Author: Bryan Maddox (University of East Anglia & CEMO, University of Oslo, Norway)
1. Utilizing process data to examine gender differences in literacy in PISA and PIAAC.

Authors
Artur Pokropek (Institute of Philosophy and Sociology of the Polish Academy of Sciences, Poland), Francesca Borgonovi (University College London, UK, and OECD, France), et Lale Khorrampel (Boston College, USA)

Presenter
Artur Pokropek (Institute of Philosophy and Sociology of the Polish Academy of Sciences, Poland)

Abstract
Gender differences in educational assessments that use standardized tests have been the subject of numerous studies for at least half a century. Strong evidence exists for two patterns: boys having higher test scores in mathematics, and girls having higher test scores in reading and literacy. To better examine the impact of problem-solving and test-taking strategies as well as differences in motivation and test-taking effort on gender differences, the current research aims to go beyond the use of classical response and rating data and will utilize log file and process data from computer-based test administrations. The proposed research aims to utilize log and process data from PISA and PIAAC cycle 1 to examine whether female and male examinees in student and adult populations use different problem solving and test-taking strategies or show differences in motivation and test-taking effort which could explain the observed differences in reading and mathematics. In a first step, existing process data indicators which are provided in the public-use data files – item-level response times, number of actions, omitted responses – will be compared between girls and boys for different item types (multiple choice versus constructed response) and at the scale and subscale level for both reading and mathematics. In a second step, big data analytics processes and data science approaches will be applied to the log data for generating new data features and indicators including data sequences, clusters, and graphs. This will be done using bottom-up as well as top-down processes and a combination of both. Theory driven decisions and analysis will be combined with expert ratings of the content domain and item development experts. Moreover, the analysis will build on findings from prior studies and modeling approaches of test-taking efforts and motivation including findings on rapid guessing behavior, careless responding and item revisit behavior.
2. Identifying strategies of information processing in PIAAC web search environments.

Authors
Carolin Hahnel, Frank Goldhammer, & Ulf Kroehne (DIPF | Leibniz Institute for Research and Information in Education, Germany)

Presenter
Carolin Hahnel (DIPF | Leibniz Institute for Research and Information in Education, Germany)

Abstract
Nowadays, it is a common activity to use online search engines to find information needed for a specific task or to solve an information problem. The variety of information provided usually enables web users to locate information quickly and in a focused way, but it also requires them to make numerous decisions (i.e., is a search entry relevant? Is the information source reliable and authoritative? Is the information sufficient to solve the information problem?). In order to make efficient use of limited time and cognitive resources, web users often employ cognitive heuristics to facilitate necessary decision processes (e.g., by choosing a search result for familiarity reasons or based on its position in the search result list).

The present study aims at investigating adults’ approaches to process web information from search engines and their success in finding an optimal solution to a given information problem. PIAAC enables this investigation by means of two items from the domain of problem-solving in technology-rich environments (PSTRE). The items required PIAAC participants to evaluate a given search engine results page and its linked websites in order to select the search entry that is optimal for solving a respective task. More specifically, the participants were requested to select the website that provided the most credible information about a medical treatment (“Sprained ankle”), or the website that offered a product that met specific criteria (“Digital Photography Book”). Participants of the PSTRE assessment were asked to solve either one or both items. Several log data indicators (i.e., time on the search result page until first website visit, number of websites visited, order of website visits) are examined using latent class analysis to identify different groups of processing behavior (e.g. groups indicating comprehensive processing of the search result page, processing focused on the comparison of websites, or disengaged test-taking). Furthermore, we examine whether the identified processing groups differ in terms of item success and, in case of adults completing both web search items, whether processing behaviors are consistent across items. The results of this log data analysis will be presented.

Authors
Qiwei He (Educational Testing Service, USA), Dandan Liao (Cambium Assessment, Inc, USA), Hok Kan Ling (Queen’s University, Canada), & Hong Jiao (University of Maryland, USA)

Presenter
Qiwei He (Educational Testing Service, USA)

Abstract
The digital assessment platform provides the possibility to collect process data such as action sequences and response times along with the responses to each task. Considering the complexity and high dimensional structure of process data, most studies that draw this new data source focus on one single item. Evaluating the behavioral consistency across items, however, renders possible capturing and modeling person-related latent characteristics, in addition to the measurement by response accuracy. This study draws on process data from log files recorded in the problem-solving in technology-rich environment (PSTRE) domain in PIAAC to evaluate the consistency of behavioral patterns across multiple items. Specifically, we presented two approaches to assess respondents’ behavioral consistency across items: (1) by using aggregate-level response process variables: the number of actions and the total response time to categorize respondents into groups based on their most likely behavioral patterns, and (2) by adding finer-grained information extracted from action sequences with sequence mining techniques. The purpose of this empirical study is twofold: first, to investigate whether the consistent behavioral patterns could be identified by process data features, and second, to examine the association among the consistency of behavioral patterns with cognitive competency and background variables. In the study sample, around 80% of respondents showed consistent patterns by the two dimensions, the response time and the number of actions, when solving multiple items. Respondents who consistently used long action sequences with short response time were found using the most similar sequences to the predefined optimal solutions. These respondents obtained the highest average PSTRE scores, had the highest ICT skills at home and at work and were the youngest group compared with their peers. Comparatively, respondents who consistently used short action sequences with fast speed had the lowest PSTRE scores and were identified as those who executed the most skipping behaviors. The findings in this study can be useful to provide information for test developers, psychometricians, and instructors for a better understanding of respondents’ consistent behavior during the cognitive process and may eventually contribute to improve task and assessment design.
4. Testing situations at home.

Author (Presenter)
Bryan Maddox (University of East Anglia & CEMO, University of Oslo, Norway)

Abstract
This paper will discuss the test quality and validity challenges associated with conducting PIAAC assessments in a household environment, and how they can be captured by 'process data'. The paper will begin by describing the ecological challenges of testing situations at home through presentation of video ethnographic case study examples on PIAAC (Maddox, 2017; Maddox & Zumbo, 2017; Maddox, 2018; Maddox, Keslair & Javr, 2019). Those highlight the scope for disturbances and assistance from by-standers – family and friends, and other distractions that can legitimately impact on the test taker experience, i.e., threats to test validity that are associated with correct test administration. The paper will describe the multiple roles of the PIAAC test administrator, in the recruitment of participants, as invigilator (proctor) in the management of the testing situation to ensure quality standards in what are sometimes unruly and idiosyncratic household interactions, and the post assessment reporting of ecological threats to validity. Informed by Latour's work in Science and Technology Studies, paper will then consider what may happen if the roles of a human administrator are replaced by 'non-human' test administration and remote invigilation. How effectively are those roles transferred to machines, and how might they be transformed in the process? The paper will consider how ecological sources of variation in the administration of test taking at home can be captured and interpreted as 'process data' (Zehner and Goldhammer, 2017; Goldhammer, Hahnel, Kroehne & Zehner, 2021) including the capture and interpretation of data from remote invigilation video data, log data on response times and keystrokes, and post assessment questionnaires.