

The Challenge of Meeting International Data Collection Standards within National Constraints: Some Examples from the Fieldwork for PIAAC in Germany

Anouk Zabal

GESIS – Leibniz Institute for the Social Sciences

Abstract

The *Programme for the International Assessment of Adult Competencies* (PIAAC) is an international OECD study that compares key competencies of adults (16-65 years) in the participating countries. In order to obtain high quality data and to ensure equivalence of measurement across countries, the international PIAAC Consortium produced a very detailed and elaborate set of standards and guidelines for all aspects of the national implementations. In Germany, a comprehensive set of measures and procedures was put in place for the PIAAC fieldwork. Some of the international requirements for data collection were not meaningful within the national context and required certain adaptations. This article describes various key fieldwork measures in Germany and discusses how specific measures relate to central international data collection standards. Reflecting on this national experience, some of the possibilities and limitations of national compliance to international standards are discussed.

Keywords: PIAAC Germany, survey standards, data collection, survey operations, fieldwork



© The Author(s) 2014. This is an Open Access article distributed under the terms of the Creative Commons Attribution 3.0 License. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1 Introduction

As a part of the *Programme for the International Assessment of Adult Competencies* (PIAAC), which was initiated and developed by the *Organization for Economic Cooperation and Development* (OECD), a first round of the PIAAC survey was carried out in 24 countries between 2008 and 2013 (OECD, 2013a).¹ The PIAAC survey continued, expanded, and refined the foundations established by two previous international large-scale assessments of adult skills: the International Adult Literacy Survey (IALS, 1994-1998; OECD & Statistics Canada, 2000) and the Adult Literacy and Lifeskills survey (ALL, 2002-2008; OECD & Statistics Canada, 2011; Statistics Canada & OECD, 2005). These further developments included extending the coverage of constructs and assessment domains, and improving the survey methodology. PIAAC strove for excellence at all stages of the survey life cycle and set very ambitious goals for the national implementations and the overall data quality. Towards this aim, the international Consortium responsible for the coordination of the PIAAC survey produced a pre-specified design, a strict timetable, and established a comprehensive program for quality assurance and quality control. As a part of the quality assurance system, an elaborate set of technical standards and guidelines (OECD, 2010) was produced to ensure that appropriate methodologies and rigorous standards be followed by all participating countries. In addition, the international Consortium closely monitored countries' work and their adherence to these technical standards and guidelines. Based on the evaluation of countries' compliance to crucial standards, a final assessment of the *fitness for use* (see Juran & Gryna, 1970; Lyberg & Biemer, 2008) of the PIAAC data for their intended purpose was undertaken by the international Consortium, together with the PIAAC Scientific Advisory Board and the Board of Participating Countries.

The countries participating in PIAAC differ significantly with regard to their type of survey organization (e.g., public organizations such as statistical agencies vs. private survey organizations), national survey practices, available sampling frames, funding, legislation, etc. Because PIAAC is a cross-national survey, the challenge thus lies in defining international survey standards that strike the appropriate balance between enforcing an adequate degree of standardization required

1 The PIAAC survey is sometimes also referred to as the Survey of Adult Skills (OECD, 2013a, 2013b).

Direct correspondence to

Anouk Zabal, GESIS – Leibniz Institute for the Social Sciences,
PO Box 12 21 55, 68072 Mannheim, Germany
E-mail: anouk.zabal@gesis.org

Acknowledgment: I would like to thank two anonymous reviewers for their constructive comments and suggestions.

for cross-national comparability, while allowing for enough degrees of freedom to accommodate differences between countries (see Koch, Blom, Stoop, & Kappelhof, 2009).

For the participating countries, translating and adapting international survey operation standards into a smoothly functioning, well-balanced, and coherent set of appropriate national measures was one of the major challenges of fieldwork. When international standards differ considerably from usual national practices, changing from tried-and-tested procedures to new, internationally prescribed ones can be risky. In extreme situations, it may even jeopardize national survey operations. That being said, novel measures and procedures can lead to innovation in national survey methods and impact positively on national survey cultures and best practices.

This article discusses efforts invested in the quality assurance and the quality control of survey operations for the main PIAAC data collection at the international level, and focuses on how some of these international standards were realized and elaborated upon, at the national level, for PIAAC in Germany. I will describe several of the key international data collection standards in PIAAC, provide an overview of the comprehensive set of measures and procedures implemented for the German fieldwork, and consider the possibilities and limitations of national compliance to international standards.

2 The PIAAC Data Collection Standards

Three chapters of the international standards and guidelines for PIAAC pertain to fieldwork survey operations (OECD, 2010): (1) field management standards, (2) data collection staff training standards, and (3) data collection standards. They address the selection, organization, and, in particular, the in-person training of the data collection staff, the data collection itself, including contact procedures, and the monitoring and quality control of fieldwork. The standards generally reflect best practices in survey operations, and the guidelines elaborate on the implementation of these standards. The approximately 65 standards and 120 guidelines, in addition to the further recommendations specified for survey operations in these chapters, go significantly beyond the breadth and depth of standards and procedures established for the precursor surveys IALS (Murray, Kirsch, & Jenkins, 1998; OECD & Statistics Canada, 2000) and ALL (OECD & Statistics Canada, 2011; Statistics Canada & OECD, 2005). The academically based, methodologically rigorous European Social Survey (ESS) also has a comprehensive set of specifications for fieldwork (e.g., European Social Survey, 2011, 2013). Many of the PIAAC standards and the ESS specifications overlap. However, the PIAAC standards and guidelines are even more elaborate than the specifications of the ESS.

Minimizing nonresponse error and increasing the rate of survey participation is at the heart of any quality survey design (see Groves & Couper, 1998). Not surprisingly, one of the most central and challenging standards in the PIAAC standards and guidelines was related to the response rates (OECD, 2010): Countries were required to achieve at least 70% overall response; however, the standards also indicated that a 50% response rate or higher would also be acceptable if the results of subsequent nonresponse bias analyses showed no evidence of significant bias. In addition, the PIAAC standards targeted a maximum non-contact rate of 3%. Both the targeted response rate and the maximum non-contact rate are the same as in the ESS (European Social Survey, 2011, 2013). In PIAAC, a data adjudication process evaluated the quality of each national data set and determined whether any limitations on the release of the data or in the international reporting should be put into effect (OECD, 2013c, Appendix 7). Response rate standards were an important element in this evaluation process. In contrast, in the ESS, although deviations from response rate standards are documented, not achieving the prescribed response rates does not have direct repercussions for the national data releases.

Many of the PIAAC standards and guidelines for fieldwork operations and data collection represent best practice methods and procedures to be implemented as a comprehensive strategy towards reaching this golden goal for response rates and to minimize nonresponse bias. Furthermore, they aim at reducing the measurement error and achieving the overall goal of collecting high-quality, internationally comparable data. Key international PIAAC standards for survey operations specified by OECD (2010) include, for example:

- (a) close monitoring of data collection at all stages,
- (b) attractive remuneration of interviewers that is independent of the number of completed interviews,
- (c) extensive in-person interviewer training,
- (d) at least four in-person contact attempts before coding a case as a non-contact,²
- (e) thorough documentation of contacting attempts and results,
- (f) no substitution of selected individuals whatsoever; use of interpreters/translators acceptable for the administration of the background questionnaire (not, however, for the cognitive assessment),
- (g) standardized administration of the survey instruments on laptops complying to specific hardware and software specifications,
- (h) development of a national best practice strategy to maximize response rates,
- (i) implementation of effective refusal conversion strategies, and

2 This is the standard for countries initially contacting the sample persons in person, which is recommended.

- (j) verification of at least 10% of each interviewer's work (random selection of all dispositions, including cases of nonresponse).³

The PIAAC standards and guidelines, extensive further documentation and material, and in-person training sessions for the National Centers, were crucial elements of the PIAAC quality assurance plan for the data collection. Compliance to key international standards was closely monitored by the international Consortium. Any proposed deviations from these standards required approval by the international Consortium. As a part of the quality control process, countries were required to fill out numerous forms and to provide information at regular intervals to keep the international Consortium updated about all aspects of national implementation and progress.

3 Key Facts about the PIAAC Data Collection

As described in more detail in OECD (2010, 2013a, 2013b, 2013c), the PIAAC interview consisted of a background questionnaire administered as a CAPI (*computer-assisted personal interview*) followed by a cognitive assessment (per default with a computer-based administration, but with the option of a paper-based administration, if required). All participating countries carried out the PIAAC interview face-to-face. In general, the interview took place at the respondent's home and was designed to take approximately 90 minutes. It was administered in the national language(s). For the background questionnaire, it was possible to recruit an interpreter to translate the questions.⁴ For the assessment, absolutely no help was allowed. Respondents worked on the cognitive assessment tasks on their own and without any time limitations. The cognitive assessment represented a non-standard requirement for both interviewers and respondents. The target population consisted of adults between 16 and 65 years of age who were non-institutionalized and were living in the country at the time of the data collection period.⁵ Countries needed to realize a probability-based sample representative of the target population. Substitutions of sampling units were not permitted at any stage.

Germany participated in the first round of the PIAAC survey, and the national implementation of the PIAAC survey was the responsibility of the German National Center at GESIS – Leibniz Institute for the Social Sciences (Rammstedt, 2013).⁶

3 In addition, 100% validation of any interviewer whose work was suspect was required.

4 The interpreter could be a family member, for example.

5 The target population was defined irrespective of nationality, residential status, or language skills.

6 The German National Center was appointed and funded by the Federal Ministry of Education and Research with the participation of the Federal Ministry of Labor and Social Affairs.

PIAAC in Germany included all domains of the cognitive assessment, i.e., literacy and numeracy, as well as the international options problem solving in technology-rich environments and reading components. Thus, the required sample size consisted of at least 5,000 cases. As indicated in the national technical report (Zabal et al., 2014), which gives a comprehensive account of all aspects of the German implementation, a registry-based sample with a two-stage stratified and clustered sampling design was realized, with 320 sample points (in 277 municipalities) and a gross sample size of 10,240 target persons. The eight-month data collection period started on 1 August 2011 and terminated on 31 March 2012. Following the PIAAC definition for a completed case (OECD, 2010), a realized sample size of 5,465 respondents was achieved in Germany. The official design-weighted final response rate for Germany (according to the PIAAC response rate definition) was 55% (Mohadjer, Krenzke, & Van de Kerckhove, 2013).⁷

4 Overview of the Fieldwork Measures in PIAAC Germany

In Germany, the data collection was subcontracted to TNS Infratest, a renowned survey organization with extensive experience in conducting face-to-face national probability-based surveys to high standards. Careful thought went into specifying a set of best practice standards and procedures for the data collection that would optimize national fieldwork and adhere as closely as possible to the PIAAC standards and guidelines, to ensure comparability and equivalence across the PIAAC countries. In order to enforce compliance with the PIAAC standards and guidelines, these were included as an appendix to the contract with the survey organization, thus emphasizing that the PIAAC data collection would entail departures from routine procedures. However, the implementation of new methods and procedures needed to be feasible within the survey organization's general organizational structure and working framework. Although the PIAAC specifications and recommendations coincided, in many instances, with best practice in Germany and in the survey organization, there were other instances where adaptations, compromises, and innovations in implementation were required.

Figure 1 shows the key elements of the German fieldwork measures. A number of these fieldwork measures are common practice for high-quality national surveys, although in general, not all measures are realized in one survey. The outstanding characteristic of the PIAAC fieldwork in Germany is that it unites a large number of measures, these measures were often undertaken with unusual intensity, and some novel methods were introduced.

⁷ The non-contact rate for Germany was 3.4% and thus only slightly above the required standard (Zabal et al., 2014).

<p>Interviewers</p> <ul style="list-style-type: none"> ▪ <i>129 experienced interviewers with excellent track records</i> ▪ <i>Five-day interviewer training</i> ▪ <i>Assigned exclusively to PIAAC for four weeks</i> ▪ <i>Attractive interviewer remuneration, including an add-on for large cities</i> 	<p>Quality Control</p> <ul style="list-style-type: none"> ▪ Thorough fieldwork monitoring by both survey organization and German National Center ▪ <i>Extended interview validation</i> ▪ <i>Monitoring of field performance (e.g. audio tapes)</i>
<p>Contacting and Gaining Cooperation</p> <ul style="list-style-type: none"> ▪ First contact in-person ▪ Four contact attempts minimum ▪ Documentation of contact attempts and further information ▪ Refusal conversion ▪ <i>Tracing respondents who had moved</i> 	<p>Incentives</p> <ul style="list-style-type: none"> ▪ Attractive conditional incentive of 50 € ▪ Small non-monetary unconditional incentive ▪ <i>Discretionary at-the-door non-monetary incentives for refusal conversion phase</i>
<p>Introductory and At-the-Door Study Materials</p> <ul style="list-style-type: none"> ▪ Advance letter ▪ <i>Brochure and flyer</i> ▪ <i>Endorsement letter and tailored letters for refusal conversion phase</i> ▪ <i>Folder with press clippings</i> 	<p>Public Relations</p> <ul style="list-style-type: none"> ▪ Press releases and <i>targeted PR work</i> ▪ Study website ▪ Toll-free hotline for respondents

Note. New measures or procedures that, in some way, went beyond standard national practice are shown in italics.

Figure 1 Key elements of the fieldwork measures in PIAAC Germany.

Fieldwork was subdivided into two main working phases and five re-issue phases (see Zabal et al., 2014). Continuous and meticulous monitoring of interviewers' work is an important aspect of survey quality control and crucial to reducing interviewer error.⁸ In PIAAC Germany, monitoring took place at various levels: (a) checking that assignments were being worked on as required (e.g., checking individual response, non-contact, and refusal rates), (b) checking the quality of the interview administration (e.g., reviewing the survey data, reviewing audio tapes), (c) validating the interviews (checking for falsifications), and (d) checking the demographic composition of the realized sample and monitoring nonresponse bias. At

⁸ This requires ongoing collection of information on interviewers' performance, the evaluation of this information, and providing interviewers with prompt feedback (see Fowler & Mangione, 1990).

the German survey organization, eight supervisors were responsible for the day-to-day operational tasks (e.g., case assignments and re-assignments, communications and instructions regarding fieldwork procedures, and feedback) and they closely monitored the interviewers. The survey organization also provided regular and detailed updates to the German National Center, which carried out further quality control and monitoring. The German National Center and the survey organization worked together closely during the entire duration of the fieldwork. In addition, regular monitoring reports were provided to the international Consortium. Issues identified during monitoring were promptly addressed with the required corrective actions.

The next sections will focus on the following subset of the key fieldwork measures listed in Figure 1: Interviewer selection, interviewer remuneration, interviewer training, incentives, contacting and gaining cooperation, and interview validation. Where appropriate, the discussion addresses tensions between the international PIAAC standards and national survey operations. A more comprehensive account of the PIAAC fieldwork, including some of the survey materials and fieldwork results, is provided in Zabal et al. (2014).

Interviewer Selection

Interviewers implement the survey design directly in their contacts with the respondents and are crucial to the quality of the survey data. With regard to the selection of the data collection staff, the PIAAC standards and guidelines recognized that numerous country-specific factors influenced the recruitment and required number of interviewers, and therefore strict standards were not prescribed. Instead, a number of considerations that countries needed to take into account were noted (Montalvan & Lemay, 2013a).

The German survey organization had a large pool of freelance interviewers at its disposal. Only experienced face-to-face interviewers with an excellent track record in the administration of high-quality registry-based CAPI surveys were considered for selection for the PIAAC survey. In addition to their experience with interview administration, interview protocols, record-keeping, and organizing their own work, these interviewers had strong interpersonal and communication skills. Experienced interviewers are more likely to be successful in gaining respondent cooperation (Groves & Couper, 1998). The selection process also took interviewers' availability for training and their availability during the eight-month fieldwork period into account (interviewers had to be able to handle their assigned workload reliably). The geographical location of the interviewers, i.e., their proximity to sample points, was also a selection factor. Only local interviewers were recruited, to maximize the number of call attempts made per case while reducing travel costs. Furthermore, local interviewers who are familiar with the area and the local dialect

and customs may achieve higher response rates than non-local interviewers (Alcser & Clemens, 2011). Several factors were considered in establishing the number of interviewers to be selected for PIAAC. In terms of reducing interviewer effects, a large number of interviewers was desirable. However, given the five-day interviewer training and the special laptop requirements for PIAAC (which necessitated the purchase of new laptops), there were pragmatic restrictions. Thus, a total of 130 freelance interviewers was selected.⁹ Most interviewers were over 50 years of age and had more than three years' experience working for the survey organization; almost 30% had more than 10 years' tenure (for more information on interviewer characteristics, see Ackermann-Piek & Massing, in this volume).

Interviewer Remuneration

Interviewer payment schemes can vary significantly across different countries and cultures (Alcser & Clemens, 2011). As a consequence, rigid standards regarding interviewer remuneration in cross-national surveys may be quite challenging, especially since specific survey organizations are unlikely to depart from their firmly established interviewer payment practices (Alcser & Clemens, 2011; Stoop, Billiet, Koch, & Fitzgerald, 2010). There are basically two standard interviewer payment arrangements: one is based on payment per completed interview, the other on an hourly rate. The advantages of a per piece payment scheme are that it is easier to monitor and it facilitates the estimation and control of interviewer costs. Paying an hourly rate is equitable in that interview length can vary substantially. Furthermore, it provides interviewers with an incentive to invest time in chasing target persons who are hard to reach or generally more reluctant to participate in surveys, and also compensates interviewers for time spent on administrative tasks and record-keeping. Finally, it discourages interviewers from speeding through the interview and undermines interviewer satisficing strategies associated with "sloppy" work. The PIAAC standards and guidelines regarding interviewer remuneration prescribed a payment per hour. The payment was to reflect the length and complexity of the PIAAC interview and be attractive in comparison with other national surveys (OECD, 2010).

As mentioned above, interviewers work on a freelance basis for the German PIAAC survey organization, as is generally the case in Germany. Consequently, the established payment for face-to-face surveys is per piece. This is markedly different from the usual practice in the United States (whose best practices in data collection shaped several PIAAC standards and guidelines), where interviewers are generally paid an hourly rate (Rosen, Murphy, Peytchev, Riley, & Lindblad, 2011). Despite

9 Due to one case of interviewer attrition prior to the start of fieldwork, 129 interviewers actually worked on the German PIAAC survey.

the weight carried by the PIAAC standards and the importance of the PIAAC survey, such a fundamental deviation from the survey organization's standard interviewer remuneration was one aspect of fieldwork which could not be changed.

As a consequence, a unique mixed payment scheme was developed for PIAAC in Germany (see Zabal et al. 2014). It consisted of three main components: (a) an attractive base rate for each completed interview, (b) an additional payment for interviews undertaken in large municipalities, and (c) an hourly payment component for interviews that were particularly long.¹⁰ The base rate per completed interview was higher than in other comparable national surveys and took into account the length and the complexity of the PIAAC interview, as well as time demands made by contact documentation tasks. The add-ons for large municipalities were introduced as a compensation for the increased interviewer burden in urban regions. In urban areas, sample persons more frequently live in dwellings with access impediments than in rural areas, and they are also less frequently at home. Thus, the add-ons for large municipalities were intended to achieve a fair, or fairer, payment across interviewers by providing additional compensation for sample points in areas with generally lower response rates and, thus, with higher interviewer burden. The hourly component for long interviews ensured that interviewers would take the time actually needed for the interview and not "rush" through. This was a crucial component, especially given that the cognitive assessment is at the heart of the PIAAC survey, and the assessment was administered without any time restriction whatsoever: Respondents worked on the cognitive tasks at their own pace and could take as long as they liked.

Interviewer Training

In PIAAC, interviewer training was regarded as a crucial feature of cross-national survey operations and as an effective tool for improving the quality of interviewers' work. Due to the complexity of the PIAAC survey, the challenging response rate goals, the importance of the PIAAC protocols both for the administration of the background questionnaire as well as for the administration of the cognitive assessment, and also given that the interview was delivered on a novel technological platform, the PIAAC international Consortium prescribed a five-day interviewer training.¹¹ Such an extensive interviewer training was a challenging novelty in Germany, where interviewer trainings are typically much shorter, if provided at all (see Zabal et al., 2014).

10 In addition, all travel costs were reimbursed.

11 For interviewers with specific profiles (experience in PIAAC field test, experience with other surveys) somewhat reduced training loads were regarded as acceptable (Montalvan & Lemay, 2013a).

At the international level, interviewer training was provided according to a train-the-trainer model (similar to the procedures in the Survey of Health, Ageing and Retirement in Europe, SHARE; see Alcsér & Benson, 2005) that aimed at ensuring consistency of training across all participating countries and hereby optimizing the standardization of interviewer behavior and survey procedures and, ultimately, ensuring the cross-national comparability of the data. The international PIAAC Consortium trained the trainers (members of the national centers, and, if possible, field directors) as if they were the interviewers and provided countries with the full set of scripted material to be translated and adapted by national centers and subsequently used for their national trainings. For some of the material, training contents required relatively extensive national adaptation (e.g., administrative survey procedures, some aspects of the background questionnaire training), whereas, for other material, any national tailoring was strictly limited, if allowed at all (e.g., administration of the cognitive assessment). In Germany, a decision was made to depart from a decentralized training solution and to have the same trainer team instruct all training groups. Training was conducted immediately before interviewers started their fieldwork to allow them to directly apply and consolidate the procedures they had learned during training.¹²

Interviewer training generally addresses two basic aspects of interviewers' work: (a) contacting target persons and gaining cooperation, and (b) the interview administration according to survey protocols. Interviewers are required to carry out a wide variety of tasks requiring both adaptive behavior as well as the capability of adhering to standardized procedures. Adaptive behavior is essential for gaining the cooperation of the sample, whereas the measurement process itself requires the ability to follow prescribed procedures in a standardized way, although interviewers also sometimes need to adapt appropriately to certain situations during the interview (Lessler, Eyerman, & Wang, 2008). Interviewer training contributes to increased survey data quality by sensitizing interviewers to respondents' concerns and to the importance of tailoring their own responses (Groves & McGonagle, 2001), as well as by decreasing item nonresponse and increasing the amount and precision of the collected information (Billiet & Loosveldt 1988).

There is no single best way to address sample persons (e.g., Groves & Couper, 1998; Groves, Singer, & Corning, 2000). Instead of using a rote introduction, it is important for interviewers to *tailor* their behavior to specific respondent characteristics and concerns, and to apply strategies to *maintain interaction* and minimize the likelihood of evoking a *no* to the survey request (see Groves & McGonagle,

12 The majority of the German interviewers participated in the full training program for interviewers with experience in other surveys (31 hours of in-person training in five days); interviewers with PIAAC field test experience took part in a reduced training (22 hours of in-person training in three days).

2001).¹³ Accordingly, during the session on gaining respondent cooperation in the German PIAAC interviewer training, interviewers practiced recognizing respondents' concerns and how to adapt their responses to specifically address these (i.e., how to tailor their responses). This included handling PIAAC-specific concerns, such as reluctance to complete the assessment. Although only experienced interviewers were assigned to PIAAC in Germany, this session was found to be invaluable, because it offered interviewers the opportunity to exchange notes and to expand their own repertoires and awareness.

Despite the fact the interviewers had extensive CAPI experience, one of the focuses of training was to specifically review the PIAAC background questionnaire and the required interviewing protocols. Groves et al. (2004) indicated that there is some evidence suggesting that experienced interviewers are not as compliant as new interviewers in reading the questions verbatim and adhering strictly to protocols. Thus, training firmly stressed the need for standardization as an important measure in reducing interviewer variance. Although the background questionnaire was developed with a view to minimizing interviewer discretion, probing, i.e., reiterating or rephrasing a question (see Cannell, Marquis, & Laurent, 1977), may sometimes be required if respondents answer inadequately. Appropriate probing techniques were also reviewed in this session.

In other sessions, interviewers were extensively trained on the administration of the cognitive assessment, which was a novel, non-standard task for them. The role change from that in the administration of the background questionnaire was an important focus. Whereas interviewers were active during the questionnaire administration, their role during the cognitive assessment was quite different. Here, interviewers had to create a quiet and supportive atmosphere, and, with the exception of technical problems, refrain from helping the respondent in any way.

Overall, training gave interviewers extensive and well-grounded knowledge about the background of the PIAAC survey, the PIAAC procedures, all components of the PIAAC interview, and the comprehensive set of materials required for the PIAAC interview. Key PIAAC standards were carefully reviewed and all measures of quality control were described in detail, to ensure full transparency. Training also included practice sessions on how to handle the novel international software. Hands-on practice exercises were found to be a crucial component of the interviewer training. Trainers circulated throughout practice interviews to observe and evaluate how interviewers were conducting the interview and whether there were any knowledge gaps to be filled or misunderstandings to clarify. Interviewers' evaluations of the training were very positive, both in the direct evaluation after training and in hindsight, as reported during the debriefings after fieldwork.

13 Interviewer training focusing on such refusal aversion strategies has been found to have certain positive effects on cooperation rates, also in face-to-face surveys (O'Brien, Mayer, Groves, & O'Neill, 2002).

Incentives

In view of the high response rate targets and the considerable respondent burden associated with the PIAAC interview, the standards and guidelines encouraged countries to consider using incentives. All planned incentives had to be signed off by the international Consortium prior to fieldwork. Whether or not a country finally opted to use an incentive for PIAAC was left to that country's discretion; in some countries, the use of incentives was not possible due to national regulations. Although prepaid incentives can be effective in increasing survey cooperation (Singer, 2002), the use of prepaid incentives for PIAAC in Germany was rejected from the outset, because prepaid incentives were not considered to be a justifiable use of taxpayers' money and are also liable to increase mistrust in and public criticism of the survey (see Börsch-Supan, Krieger, & Schröder, 2013). Three conditional incentive conditions were tested in the German PIAAC field test, to determine the best incentive for the main survey (a 10 € commemorative coin, 25 € in cash, or 50 € in cash).¹⁴ Following evaluation of the results of the German field test incentive experiment, the largest of the tested incentives, the 50 € conditional cash incentive was chosen for the main survey (for details, see Martin, Helmschrott, & Rammstedt, in this volume). In addition, a non-monetary unconditional incentive (post-it notes featuring the PIAAC logo) was attached to the advance letter. In the re-issue phases, interviewers were given the option of deploying discretionary at-the-door non-monetary incentives. The 50 € incentive constitutes a substantive sum, in comparison to incentives offered by other national surveys (see Pffor et al., forthcoming). It reflects not only the national importance of the PIAAC survey, but also the substantial length of the interview, and acknowledges that participating in an assessment is an unusual and, for some, daunting aspect of the interview.

Contacting and Gaining Cooperation

Before any interview can be carried out, the interviewer has to locate the target person, establish contact, and gain their cooperation. The majority of the PIAAC standards and guidelines relating to contacting and callback rules, study materials and outreach tools, as well as techniques for dealing with nonresponse cases were in line with many national best practices in Germany. The contacting rules for PIAAC in Germany ascertained that at least four in-person contact attempts be made before a non-contact could be coded, with calls to take place at different times of the day and on different days of the week, to accommodate varying at-home patterns and facilitate reaching difficult-to-contact sample persons. In

14 All countries participating in Round 1 of PIAAC conducted a field test in 2010. Some information on the German PIAAC field test is provided in Zabal et al. (2014).

addition, interviewers were required to record each contact attempt and the disposition of each contact outcome. Prior to the first contact attempt, an advance letter, accompanied by a study flyer and the unconditional incentive, was sent to the sample person. Contrary to the survey organization's usual practice, a staggered mailing schedule was implemented that was individually attuned to each interviewer's personal contacting schedule, in an effort to reduce the time interval between receipt of the advance letter and the interviewer's first visit. Another new measure at the survey organization consisted in assigning interviewers exclusively to the PIAAC survey for four weeks during the first phase of fieldwork. Furthermore, the German National Center undertook considerable efforts to produce attractive study materials (e.g., not only a flyer but also a brochure), and, in targeted public relations activities, to increase the visibility of the PIAAC survey in Germany and emphasize the legitimacy of the interviewer's request (study website, toll-free hotline for respondents, press releases, and the targeted dispatch of the press releases to local newspapers in the PIAAC sample points).

During the re-issue phases, only a subset of the refusals could be re-approached, due to German legislation; "hard refusals" could not be re-contacted. For those cases that could be re-issued, additional refusal conversion measures were introduced: (a) tailored refusal conversion letters reinforcing specific aspects of the survey, (b) extended at-the-doorstep material that included an endorsement letter and translations of the advance letter and FAQs into five languages, (c) discretionary at-the-door non-monetary incentives, (d) re-assignment of interviewers, and (e) a selective deployment of travelling interviewers to difficult sample points.

One of the issues unique to countries with registry samples is that the selected addresses may be obsolete by the time a contact with the sample person is attempted. For example, persons who had re-located may not have correctly deregistered and re-registered. This problem is exacerbated in countries such as Germany that do not have a central population register but have nationally distributed registry offices. In registry-based high-quality surveys in Germany, it is common practice to classify cases with address-related dispositions as ineligible (e.g., in the German General Social Survey, ALLBUS; see Wasmer, Scholz, Blohm, Walter, & Jutz, 2012). However, this was not an option in the context of the PIAAC standards. Instead, it was necessary to undertake special efforts to trace respondents who had moved or whose addresses proved to be invalid. To cope with this situation, a new procedure was introduced: cases with unresolved address-related dispositions¹⁵ were collected at home office, and the registry offices were subsequently re-contacted with a request for updated information. This approach proved to be useful; for details, see Zabal et al. (2014). In addition, and contrary to common practice, respondents

15 Resolved address-related dispositions were cases in which the sample person had moved outside of the country or for which the interviewer was successful in obtaining a new address.

who had moved to non-PIAAC sample points were also pursued (within certain feasibility limits).

Interview Validation

Interviewer falsification denotes intentional interviewer deviations from the survey protocols, such as the fabrication of interviews (or parts thereof), the substitution of sample persons, misreporting disposition codes, or taking shortcuts through the interview (see Groves et al., 2004). One of the most important PIAAC standards with regard to identifying possible falsifications stated that 10% of each interviewer's finalized work had to be verified, including final nonresponse dispositions (OECD, 2010). In addition, one of the guidelines for this standard stipulated that cases for verification should be selected at random from all sampling units (including both respondents and nonrespondents). For PIAAC Germany, this standard and guideline were fundamentally problematic. The survey organization's common validation practice is to validate *all completed interviews*, and only these. This strategy is based on two considerations. First, because one of the potential drawbacks associated with the usual per piece payment is a higher risk of interviewer falsification (Rosen et al., 2011), the focus is clearly on identifying any potentially falsified interviews. Second, German legislation prohibits re-approaching hard refusals. After intensive deliberation, it was decided that departing from the survey organization's well-established validation procedures posed too great a risk. Thus, its standard validation procedure was adopted as a starting point. It essentially consists of sending all respondents a validation questionnaire by mail, and in exploiting one of the advantages of using a registry-based sample by checking the consistency of interview data with the basic information provided by the register when the sample is drawn (age, gender, and nationality). Furthermore, it includes a number of additional checks (e.g., interview time and length). Conspicuous cases are systematically followed up.

For PIAAC, the back-checks were extended to include other dispositions – as far as feasible, but with definite limitations, e.g., hard refusals could not be validated by law. Concretely, attempts were made to validate:

- (a) certain ineligibles: via an internet search (ineligibles due to institutionalization),
- (b) refusals due to disabilities: via a mail validation questionnaire,
- (c) non-contacts: through a concerted telephone action, and
- (d) soft refusals: in person.

Although the standard validation procedures worked smoothly and ensured that at least 10% of each interviewer's work was successfully validated, the attempted extensions of the validation scheme yielded only very modest returns. Checking

the ineligible due to institutionalization proved to be practicable. The back-checks of refusals due to disabilities were not found to be advisable, due to certain ensuing ethical issues. The attempt to reach and validate non-contacts by phone was relatively unsuccessful, and the procedure for validating a non-contact was generally debatable. The main focus in re-approaching soft refusals remained refusal conversion (and not validation). Furthermore, the attempt to validate soft refusals in person during the refusal conversion phase did not work as smoothly as intended. With respect to the random selection specification, because all completed cases and all disabilities were selected for validation, the complete selection was superior to selecting a random subset. Given the legal restrictions in re-approaching hard refusals, a random selection from all dispositions was not possible in Germany.

Beyond carefully reviewing the survey organization's quality control results, the German National Center carried out a set of validation measures that complemented the basic validation described above. These included (a) monitoring the date and time of the interview and number of interviews per day, (b) monitoring the length of the interviews to identify suspicious outliers, especially scrutinizing very short interviews, (c) checking some interviews for routing shortcuts, (d) reviewing item nonresponse rates, (e) reviewing the quality of the entered responses to certain open format questions, and (f) checking the quality of the interviewers' scoring.¹⁶

Reviewing audio tapes of actual interviews provides direct information on the interview process (Fowler & Mangione, 1990). The PIAAC standards specified that each interviewer had to submit two tape recordings of interviews early on during fieldwork, with subsequent review of the recordings (OECD, 2010). This specification addressed the need to check if the PIAAC protocols and procedures taught in training were being applied appropriately. However, it should be noted that the use of tape recorders may lead interviewers to perform better (Billiet & Loosveldt, 1988). Although it was a non-standard requirement within the German framework for fieldwork, the vast majority of interviewers did, in fact, deliver audio tapes for monitoring. These audio tapes (specifically, the recording of the background questionnaire administration) were systematically reviewed at the German National Center. If deviations from the protocols were found, for example incorrect use of show cards or a tendency to not read each question fully and accurately, the survey organization was contacted with instructions to re-train specific interviewers on the identified issues. Ackermann-Piek and Massing (in this volume) describe these audio tape reviews in more detail and provide some evaluations of the interviewers' behavior.

16 Scoring denotes the evaluation of responses to cognitive tasks and coding them as correct or incorrect. One of the more difficult and training-intensive PIAAC interviewer tasks involved scoring responses to eight core assessment tasks that were part of the paper-based assessment.

Interview validation inspected the overall patterns of all these measures and closed followed-up on any conspicuous constellations. In Germany, no instance of falsification was detected. Further information on the interview validation and fieldwork quality control in PIAAC Germany can be found in Massing, Ackermann, Martin, Zabal, and Rammstedt (2013) and in Zabal et al. (2014).

5 Discussion

Section 4 describes key parameters of the German fieldwork strategy for PIAAC. These included the best possible and most comprehensive set of fieldwork measures that would work well within the national context, within the context of the national survey organization, and within the framework established by the PIAAC standards and guidelines. These procedures were followed rigorously during all phases of data collection, to obtain results of the highest possible quality. Overall, the German fieldwork strategy worked well and was effective in reaching national and international data collection goals. In the light of the general decline in response rates for face-to-face surveys in Germany and many other countries (e.g., Blohm & Koch, 2013; de Leeuw & de Heer, 2002), the achieved weighted response rate of 55% for PIAAC in Germany can be regarded as a particularly successful outcome.

Groves and Couper (1998) describe survey participation as a function of several factors that are grounded in features of the survey design, environmental features, individual characteristics of the sample persons, as well as in characteristics of the interviewer and the interviewer-sample person interaction. Thus, a wide variety of factors may affect a sample person's decision to participate in a survey, ranging from the survey climate, and the trustworthiness and respectability of the sponsor, to the subjective burden associated with the survey request, or the appeal of the offered incentive. Some of the international design specifications of the PIAAC survey that were beyond the control of the national implementation are potentially detrimental to gaining cooperation; for example, the interview length, and, at least to some extent, the request to participate in a cognitive assessment. On the other hand, a number of other factors are especially favorable, such as the long data collection period, and, in particular, the prominence of PIAAC, its international dimension, and its political relevance.

For Germany, PIAAC was a survey of particular national importance. This was a decisive factor that impacted on the national implementation, both directly and indirectly. Due to the priority of PIAAC, it was well funded and, as a consequence, the range of possible measures and interventions was larger than usual. This was an important element in realizing the sophisticated fieldwork strategy required to achieve internationally comparable and high-quality survey results. It also made it possible to offer an unusually attractive incentive of 50 €. Further-

more, the survey organization clearly acknowledged its internal prioritization of the PIAAC survey. It was therefore possible to initiate more novel components and modifications to standard procedures than usual. It should be noted that including the comprehensive PIAAC standards and guidelines as a part of the contract with the survey organization seems to have significantly contributed to triggering improved methods and departing from routine practice.

As previously indicated, while many of the PIAAC standards reflect national best practice, others do not. At the onset of the PIAAC survey, there were certain misgivings about the feasibility of a number of these standards in Germany. In some cases, these reservations proved to be wrong. For example, both the necessity and the feasibility of a five-day interviewer training were questioned. However, it turned out to be both necessary and feasible. The five days were indeed needed to review and transmit all the relevant information regarding the complex PIAAC interview, and to ensure that interviewers could smoothly bring together all the various components and procedures. The length of training was also justified by the need for standardization across the participating countries, by the introduction of diverse novel elements to the interviewers' work, and by the deviations from their usual practice. The latter should not be underestimated: Lynn (2003, p. 330) emphasizes that "The potential for errors and mistakes when people used to doing things one way are asked to do them in a slightly different way is considerable." Beyond these objective reasons, the interviewer training in Germany was found to have unforeseen and very positive motivational side-effects for interviewers and home office staff. Spending five intensive days together contributed to a sincere team building between all players – interviewers, supervisors, field directors, members of the German National Center – and created a strong identification with the PIAAC survey and its aims. To sum, in hindsight the interviewer training was vital for fieldwork success. However, this is not an appeal to widely implement five-day interviewer trainings for all German surveys. Many surveys will have neither a pressing need (in terms of the complexity of the interview and protocols), nor the resources for such (extended) in-person trainings. If extensive in-person trainings became a standard, they would also no longer have the unique motivational effect that they had for PIAAC. However, it is important to emphasize that even experienced interviewers can profit greatly from training on gaining cooperation and on standardized interviewing techniques.

Another example of a standard that was first thought to be problematic in the German context was the requirement to obtain audio tapes of actual interviews. Contrary to the initial forecasts, being asked to audio tape an interview did not cause significant friction, neither with the interviewer, nor the respondents. Admittedly, it remains unclear how well this would have been received without the PIAAC training, the weight of the international PIAAC standards, and the attractive conditional incentive. From the point of view of quality control, this direct monitoring

is especially suited to identifying interviewer mistakes in administering the interview. This aspect of quality control is not pursued in many national surveys, and experience with PIAAC shows that even very experienced interviewers can deviate from standardized survey protocols, and that monitoring the CAPI administration and providing timely feedback is important. It would therefore be recommendable to consider adopting this quality control element in other national surveys. However, it should be noted that reviewing the audio tapes required significant personnel resources at the German National Center, and that not all surveys will have the capacity needed for this work.

Some of the PIAAC data collection standards remained unfeasible in the German context, despite endeavors to achieve compliance. For example, the central component of the national interviewer remuneration remained a per piece and not a per hour payment. However, the national extensions to the standard remuneration practice captured the spirit of the international standards, which, in essence, consisted in providing an attractive and equitable payment for all aspects of the interviewers' tasks. In this case, it represented a viable compromise between the international requirements and national possibilities.

The case is different for the interview validation scheme. Here, national legislation and well-established validation procedures were diametrically opposed to the international standards. Without intending to imply that the national validation strategy cannot be improved upon, it is a strategy that harmonizes with other national fieldwork elements and makes sense in the German national context. Quality control back-checks are such a crucial element of fieldwork that completely changing the validation approach for one survey is neither feasible nor recommendable. The risk involved in departing completely from well-established procedures of this importance is significant. Based on the traditional national validation approach, and with every effort made to put in practice the entire array of additional possible checks, as well as introducing completely new ones to extend the range of validated dispositions, validation in PIAAC Germany was very thorough and comprehensive. However, the attempt to match the international validation scheme more closely by implementing new quality control back-checks for non-complete dispositions did not work very well.

From an international perspective, the detailed information provided by Montalvan and Lemay (2013a) about several aspects of fieldwork operations in the PIAAC Round 1 countries presents a useful overview of variations across countries. Montalvan and Lemay (2013b) also describe the quality assurance and quality control activities for the PIAAC survey operations and conclude that countries' compliance with the quality control program was high. As mentioned above, the comprehensive quality control mechanisms put in place for PIAAC culminated in a final data adjudication process. The development of the adjudication framework and the selection of indicators were undertaken relatively late in the survey life-

cycle. OECD (2013c, Appendix 7) describes the process and results of the final data adjudication. This data adjudication went beyond the mere evaluation of compliance with PIAAC standards. It aimed at evaluating the overall quality of the PIAAC data in terms of their "fitness for use", i.e., to assess whether the quality of the data was sufficient for the intended use (e.g., to inform policy-makers, for scientific purposes), or whether restrictions needed to be imposed on the dissemination and use of the data. Data collection was one of the four core domains scrutinized in the final data adjudication process; the other domains were sampling, coverage and nonresponse bias, and instrumentation. Each domain was evaluated according to a set of indicators, with three possible outcomes (pass, caution, or fail) that reflected whether the relevant requirements were fully met, met to an acceptable extent, or generally not fulfilled. The German data collection was given a pass, with a comment that the validation strategy met a reduced requirement (OECD, 2013c, Appendix 7). The requirements regarding response rates and coverage rates were considered as a part of the data adjudication domain "coverage and nonresponse bias". For Germany, the data adjudication report noted a caution for this domain but indicated that the extended nonresponse bias analysis "provides evidence that bias was reduced through the weighting adjustments" (OECD, 2013c, Appendix 7, p. 70). It is interesting to note that while the results of the nonresponse bias analyses were clearly essential for the evaluation of this domain, only the five countries with a final weighted response rate of 70% or above were given a pass. All countries with a final weighted response rate below 70% were assigned a caution (as was the case in Germany).

6 Conclusions and Outlook

Surveys such as PIAAC that strive to achieve cross-national comparability and produce data of the highest possible quality by implementing an effective system of quality assurance and control, and that receive high priority at international and national levels, have the potential to bring about welcome innovation to national survey practices. Countries participating in PIAAC had a strong incentive to reach the exacting international standards and, as such, these standards were often the gate-openers to adapting standard methods and procedures and adopting new survey operations. There were many instances of this in the fieldwork for PIAAC in Germany. Beyond the examples discussed in the previous section, many other details of fieldwork were adjusted or improved upon for the German implementation of PIAAC. Some of these may enrich future national surveys (e.g., address search).

Standardization of survey operations aims at achieving comparability. Even though the need for standardization in the data collection of cross-national com-

parative surveys is uncontested, there are also limits to standardization. Occasionally, comparability of results is better achieved by deliberately doing some things differently (Harkness, 2008; Koch et al., 2009; Lynn, 2003). Some of the PIAAC data collection standards and guidelines already explicitly allowed for different approaches, depending on countries' typical survey procedures. Furthermore, in the process of international quality control, certain country deviations were regarded as acceptable alternatives (Montalvan & Lemay, 2013a). Some of the other data collection standards, however, made no such allowances for national variability. The experience with the German PIAAC fieldwork, most aptly illustrated by the example of interview validation, points to the need for further reflection on how best to reach cross-national comparability in survey operations. It is thus with reservation that we note the recommendation proposed by Montalvan and Lemay (2013a) for future cycles of PIAAC calling for unconditional adherence to all validation standards and guidelines, specifically, the random selection of all finalized cases at a 10% level.

The "best" survey operations will differ, depending on the specific countries (and even on the specific survey organizations) involved. The challenge in defining an appropriate set of international standards is to strike the right balance between standardization and national adaptations (see Koch et al., 2009). Because cross-national differences exist, it may not always be possible to define single standards that are realistically achievable in all countries. Furthermore, the implementation of international survey operation standards will have different repercussions in different countries, including costs and timelines (Lynn, 2003). Thus, it is advisable to involve countries in the process of setting standards to make it, at least partly, a collaborative effort, with national conditions shaping the international survey standards and determining their relative importance. Lastly, in order to achieve full transparency in the program of quality assurance, it is crucial that not only the survey standards be known at the onset of the survey, but also the relevant framework and indicators for the data adjudication.

References

- Ackermann-Piek, D., & Massing, N. (2014). Interviewer behavior and interviewer characteristics in PIAAC Germany. *methods, data, analyses*, 8(2), 199-222. doi: 10.12758/mda.2014.008
- Alcser, K. H., & Benson, G. (2005). The SHARE train-the-trainer program. In A. Börsch-Supan & H. Jürges (Eds.), *The Survey of Health, Ageing, and Retirement in Europe – Methodology* (pp. 70-74). Mannheim: MEA. Retrieved from http://www.share-project.org/uploads/tx_sharepublications/Methodology_Ch6.pdf

- Alcser, K. H., & Clemens, J. (2011). X. Interviewer recruitment, selection, and training (Revised 28. Nov. 2011). In *Cross cultural survey guidelines*. Retrieved from <http://ccsg.isr.umich.edu/pdf/10InterviewerRecruitmentFeb2012.pdf>
- Billiet, J., & Loosveldt, G. (1988). Improvement of the quality of responses to factual survey questions by interviewer training. *Public Opinion Quarterly*, 52(2), 190-211.
- Blohm, M., & Koch, A. (2013). Respondent incentives in a national face-to-face survey: Effects on outcome rates, sample composition and fieldwork efforts. *Methoden, Daten, Analysen (MDA)*, 7(1), 89-122. doi: 10.12758/mda.2013.004
- Börsch-Supan, A., Krieger, U., & Schröder, M. (2013). *Respondent incentives, interviewer training and survey participation*. SHARE Working Paper (12-2013). Munich Centre for the Economics of Ageing (MEA). Retrieved from <http://www.share-project.org/publications/workingpapers0.html>
- Cannell, C. F., Marquis, K. H., & Laurent, A. (1977). *A summary of studies of interviewing methodology*. Washington, D.C.: US Government Printing Office.
- de Leeuw, E. D., & de Heer, W. (2002). Trends in household survey nonresponse: A longitudinal and international comparison. In R. M. Groves, D. A. Dillman, J. L. Eltinge & R. J. A. Little (Eds.), *Survey nonresponse* (pp. 41-54). New York, NY: John Wiley & Sons.
- European Social Survey. (2011). *Round 6 specification for participating countries*. London: Centre for Comparative Social Surveys, City University London.
- European Social Survey. (2013). *Round 7 specification for participating countries*. London: Centre for Comparative Social Surveys, City University London.
- Fowler, F. J., & Mangione, T. W. (1990). *Standardized survey interviewing: Minimizing interviewer-related error*. Newbury Park, California: Sage Publications.
- Groves, R. M., & Couper, M. P. (1998). *Nonresponse in household interview surveys*. New York, NY: John Wiley & Sons.
- Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2004). *Survey methodology*. Hoboken, NJ: John Wiley & Sons.
- Groves, R. M., & McGonagle, K. A. (2001). A theory-guided interviewer training protocol regarding survey participation. *Journal of Official Statistics*, 17(2), 249-265.
- Groves, R. M., Singer, E., & Corning, A. (2000). Leverage-saliency theory of survey participation. *Public Opinion Quarterly*, 64, 299-308.
- Harkness, J. A. (2008). Comparative survey research: Goal and challenges. In E. D. de Leeuw, J. J. Hox & D. A. Dillman (Eds.), *International handbook of survey methodology* (pp. 56-77). New York, NY: Psychology Press.
- Juran, J. M., & Gryna, F. M. (1970). *Quality planning and analysis*. New York, NY: McGraw-Hill Inc.
- Koch, A., Blom, A. G., Stoop, I., & Kappelhof, J. (2009). Data collection quality assurance in cross-national surveys: The example of the ESS. *Methoden, Daten, Analysen (MDA)*, 3(2), 219-247.
- Lessler, J. T., Eyerman, J., & Wang, K. (2008). Interviewer training. In E. D. de Leeuw, J. J. Hox & D. A. Dillman (Eds.), *International handbook of survey methodology* (pp. 442-460). New York, NY: Psychology Press.
- Lyberg, L. E., & Biemer, P. P. (2008). Quality assurance and quality control in surveys. In E. D. de Leeuw, J. J. Hox & D. A. Dillman (Eds.), *International handbook of survey methodology* (pp. 421-441). New York, NY: Psychology Press.
- Lynn, P. (2003). Developing quality standards for cross-national survey research: Five approaches. *International Journal of Social Research Methodology*, 6(4), 323-336.

- Martin, S., Helmschrott, S., & Rammstedt, B. (2014). The use of respondent incentives in PIAAC: The field test experiment in Germany. *methods, data, analyses*, 8(2), 223-242. doi: 10.12758/mda.2014.009
- Massing, N., Ackermann, D., Martin, S., Zabal, A., & Rammstedt, B. (2013). Controlling interviewers' work in PIAAC – the Programme for the International Assessment of Adult Competencies. In P. Winker, N. Menold & R. Porst (Eds.), *Interviewers' deviations in surveys. Impact, reasons, detection and prevention* (pp. 117-130). Frankfurt am Main: Peter Lang.
- Mohadjer, L., Krenzke, T., & Van de Kerckhove, W. (2013). Indicators of the quality of the sample data. In OECD (Ed.), *Technical report of the Survey of Adult Skills (PIAAC)* (Chapter 16, pp. 1-30). Paris: OECD. Retrieved from <http://www.oecd.org/site/piaac/publications.htm>
- Montalvan, P., & Lemay, M. (2013a). Field operations. In OECD (Ed.), *Technical report of the Survey of Adult Skills (PIAAC)* (Chapter 10, pp. 1-38). Paris: OECD. Retrieved from <http://www.oecd.org/site/piaac/publications.htm>
- Montalvan, P., & Lemay, M. (2013b). Quality control monitoring activities. In OECD (Ed.), *Technical report of the Survey of Adult Skills (PIAAC)* (Chapter 11, pp. 1-12). Paris: OECD. Retrieved from <http://www.oecd.org/site/piaac/publications.htm>
- Murray, T. S., Kirsch, I. S., & Jenkins, L. B. (1998). *Adult literacy in OECD countries: Technical report on the first International Adult Literacy Survey* (NCES 98-053). Washington D.C.: U.S. Department of Education, National Center for Education Statistics.
- O'Brien, E. M., Mayer, T. S., Groves, R. M., & O'Neill, G. E. (2002). Interviewer training to increase survey participation. *Proceedings of the Survey Research Methods Section* (pp. 2502-2507). Alexandria: American Statistical Association. Retrieved from <http://www.amstat.org/sections/srms/Proceedings/y2002/Files/JSM2002-000530.pdf>
- OECD. (2010). *PIAAC technical standards and guidelines* (December 2010). Retrieved March 2014 from <http://www.oecd.org/site/piaac/surveyofadultskills.htm>
- OECD. (2013a). *OECD Skills outlook 2013: First results from the Survey of Adult Skills*. Paris: OECD Publishing.
- OECD. (2013b). *The Survey of Adult Skills – Reader's companion*. Paris: OECD Publishing.
- OECD. (2013c). *Technical report of the Survey of Adult Skills (PIAAC)*. Paris: OECD. Retrieved from <http://www.oecd.org/site/piaac/publications.htm>
- OECD & Statistics Canada. (2000). *Literacy in the information age. Final report of the International Adult Literacy Survey*. Paris: OECD Publishing.
- OECD & Statistics Canada. (2011). *Literacy for life: Further results from the Adult Literacy and Life Skills Survey*. Paris: OECD Publishing.
- Pffor, K., Blohm, M., Blom, A. G., Erdel, B., Felderer, B., Fräbendorf, M., . . . Rammstedt, B. (forthcoming). Are incentive effects on response rates and nonresponse bias in large-scale, face-to-face surveys generalizable to Germany? Evidence from ten experiments. *Public Opinion Quarterly*.
- Rammstedt, B. (2013). *Grundlegende Kompetenzen Erwachsener im internationalen Vergleich – Ergebnisse von PIAAC 2012*. Münster: Waxmann.
- Rosen, J., Murphy, J., Peytchev, A., Riley, S., & Lindblad, M. (2011). The effects of differential interviewer incentives on a field data collection effort. *Field Methods*, 23(1), 24-36.
- Singer, E. (2002). The use of incentives to reduce nonresponse in household surveys. In R. M. Groves, D. A. Dillman, J. L. Eltinge & R. J. A. Little (Eds.), *Survey Nonresponse* (pp. 163-177). New York, NY: John Wiley & Sons.

- Statistics Canada & OECD. (2005). *Learning a living. First results of the Adult Literacy and Life Skills Survey*. Paris: OECD Publishing.
- Stoop, I., Billiet, J., Koch, A., & Fitzgerald, R. (2010). *Improving survey response: Lessons learned from the European Social Survey*. Chichester: John Wiley & Sons.
- Wasmer, M., Scholz, E., Blohm, M., Walter, J., & Jutz, R. (2012). *Konzeption und Durchführung der „Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften“ (ALLBUS) 2010* (GESIS-Technical Reports 2012/12). Köln: GESIS – Leibniz-Institut für Sozialwissenschaften.
- Zabal, A., Martin, S., Massing, N., Ackermann, D., Helmschrott, S., Barkow, I., & Rammstedt, B. (2014). *PIAAC Germany 2012: Technical report*. Münster: Waxmann.