

### PIAAC-L data collection 2016: technical report

Martin, Silke; Zabal, Anouk; Rammstedt, Beatrice

Veröffentlichungsversion / Published Version

Arbeitspapier / working paper

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

#### Empfohlene Zitierung / Suggested Citation:

Martin, S., Zabal, A., & Rammstedt, B. (2018). *PIAAC-L data collection 2016: technical report*. (GESIS Papers, 2018/05). Köln: GESIS - Leibniz-Institut für Sozialwissenschaften. <https://doi.org/10.21241/ssoar.56852>

#### Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-NC Lizenz (Namensnennung-Nicht-kommerziell) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:  
<https://creativecommons.org/licenses/by-nc/4.0/deed.de>

#### Terms of use:

This document is made available under a CC BY-NC Licence (Attribution-NonCommercial). For more information see:  
<https://creativecommons.org/licenses/by-nc/4.0>

gesis

Leibniz-Institut  
für Sozialwissenschaften

GESIS Papers

2018|05

**PIAAC-L Data Collection 2016**

**Technical Report**

*Silke Martin, Anouk Zabal, &  
Beatrice Rammstedt*



GESIS Papers 2018|05

## **PIAAC-L Data Collection 2016**

### **Technical Report**

*Silke Martin, Anouk Zabal, &  
Beatrice Rammstedt*

## **GESIS Papers**

GESIS – Leibniz-Institut für Sozialwissenschaften  
Survey Design and Methodology  
Postfach 12 21 55  
68072 Mannheim  
Telefon: (0621) 1246 - 518  
Telefax: (0621) 1246 - 100  
E-Mail: [anouk.zabal@gesis.org](mailto:anouk.zabal@gesis.org)

*PIAAC-L was funded by the Federal Ministry of Education and Research, Berlin (Grant number 01 JP 1301 A, B, C). We thank the other members of the PIAAC-L team for their work in realizing the second wave of PIAAC-L.*

*Correspondence concerning this manuscript should be addressed to Silke Martin, GESIS – Leibniz Institute for the Social Sciences, PO Box 12 21 55, 68072 Mannheim, Germany. E-mail: [silke.martin@gesis.org](mailto:silke.martin@gesis.org)*

ISSN: 2364-3781 (Online)  
Publisher, printing  
and distribution: GESIS – Leibniz Institute for the Social Sciences  
Unter Sachsenhausen 6-8, 50667 Cologne, Germany

## Contents

---

1	Introduction.....	5
2	Wave 3 in PIAAC-L .....	8
2.1	Wave 3: The Final PIAAC-L Data Collection.....	8
2.2	The PIAAC-L Sample in Wave 3 .....	9
3	Instruments.....	11
3.1	Contact Documentation and Household Protocol .....	11
3.2	Household Questionnaire .....	12
3.3	Person Questionnaire .....	13
3.3.1	Core person questionnaire .....	13
3.3.2	Basic cognitive skills (short scales) .....	15
3.3.3	The Number Series Study.....	16
4	Fieldwork: Preparation, Operations, and Results .....	17
4.1	Preparatory Work.....	17
4.2	Field Operations.....	19
4.3	Fieldwork Results.....	23
5	Data Processing and Data Products.....	27
5.1	Data Cleaning, Processing, and Coding.....	27
5.2	Measurement of Basic Cognitive Skills in Wave 3.....	28
5.3	Weighting .....	29
5.4	Data Dissemination.....	29
	References.....	32
	Annex: Questions From Original SOEP Instruments.....	34

## 1 Introduction

---

This paper is part of a series of reports describing the technical implementation of PIAAC-L, the German PIAAC-Longitudinal project funded by the Federal Ministry of Education and Research. It is a follow-up to the technical reports for wave 1 (Zabal, Martin, & Rammstedt, 2016) and wave 2 (Zabal, Martin, & Rammstedt, 2017) and aims to describe the design, instruments, fieldwork processes, and data dissemination for wave 3.

PIAAC-L is a national survey that followed-up the German respondents who participated in the international PIAAC survey (*Programme for the International Assessment of Adult Competencies*) organized by the *Organisation for Economic Cooperation and Development* (OECD). With PIAAC, which was conducted in 2011/12, key adult skills in the domains literacy, numeracy, and problem solving in technology-rich environments were assessed multi-nationally in a cross-sectional design (OECD, 2013a, 2013b; Rammstedt, 2013; for information on the technical implementation in Germany see Zabal et al., 2014). With PIAAC-L a source of data is created that serves to enhance the rich body of data obtained in PIAAC, and by that allows for analyses of a variety of research questions and topics. PIAAC-L is a joint project of three partner institutes: GESIS – Leibniz Institute for the Social Sciences (lead; formerly National Center for PIAAC Germany), the Socio-Economic Panel (SOEP) at the German Institute for Economic Research (DIW Berlin), and the Leibniz Institute for Educational Trajectories (LifBi; responsible for the National Educational Panel Survey, NEPS).

PIAAC-L is designed as a three-wave follow-up survey to PIAAC respondents, with data collections in 2014, 2015, and 2016. While foci and groups of addressed persons vary somewhat across waves, PIAAC respondents, called *anchor persons*, are consistently the central response units in PIAAC-L. Wave 1 was designed to target anchor persons and their household members aged 18 and over (born 1996 or earlier), using core household and person questionnaires from the SOEP. In wave 2, anchor persons and their partners, if living in the same household, were addressed. In addition to the administration of a newly assembled person questionnaire with questions from a variety of source instruments (such as PIAAC, SOEP, NEPS), respondents took a cognitive assessment using PIAAC instruments for literacy and numeracy and/or NEPS instruments for reading and mathematics. The design of the third wave was similar to that of the first wave: Anchor persons and all household members aged 18 and over (born 1998 or earlier) were to be interviewed; a household and a person questionnaire were administered. Although some questions were taken from SOEP questionnaires, the instruments were adapted to include questions newly developed by researchers from the PIAAC Leibniz network.<sup>1</sup> In addition, two modules assessing cognitive skills were administered. Design specifications of this last wave of PIAAC-L are illustrated in the present paper. The rationale for implementing PIAAC-L and overall design information are summarized in Rammstedt, Martin, Zabal, Carstensen, and Schupp (2017).

---

<sup>1</sup> Select *PIAAC Leibniz Network (PIAAC-LN)* at <https://www.gesis.org/en/piaac/projects/>

Box 1.1 replicates the summary of general key facts for PIAAC-L from Zabal et al. (2016, p. 6). Box 1.2 summarizes the key facts for the data collection in 2016.

**Box 1.1: Key Facts: PIAAC-Longitudinal (PIAAC-L), Germany**

- National longitudinal follow-up of PIAAC Germany 2012 respondents
- Three waves of data collection: 2014, 2015, 2016
- Main objective is to extend and enhance analytical potential of German PIAAC 2012 data
- Cooperative undertaking between three major Leibniz institutes and surveys:
  - GESIS – Leibniz Institute for the Social Sciences [lead]
  - Socio-Economic Panel (SOEP) at German Institute for Economic Research (DIW Berlin)
  - Leibniz Institute for Educational Trajectories (LifBi)
- Survey organization: TNS Infratest (now: Kantar TNS)
- Funded by the Federal Ministry of Education and Research

**Box 1.2: Key Facts: The Data Collection 2016 (PIAAC-L Wave 3)**

- Instruments:
  - Household questionnaire (based on SOEP core instrument)
    - Living situation, conditions, and costs
    - Household income and benefits, wealth
    - Household possessions and customs
    - Children
  - Person questionnaire (parts from SOEP core instrument & new questions)
    - Background information, family, childhood, and cultural identity
    - Biographical calendar
    - Formal education (general and vocational education)
    - Adult education and training (vocational and general further training)
    - Work status, work situation and history, change of employer, skill mismatch
    - Income and benefits
    - Health, satisfaction, and attitudes
    - Time use and leisure activities
    - Life events
  - Basic cognitive skills assessed with short scales (from SOEP)
    - Animal Naming Task
    - Multiple-Choice Vocabulary Intelligence Test
    - Symbol-Digit Test
  - Number Series Study (add-on module by the German Institute for International Educational Research, DIPF)

- Interview administration: CAPI (computer-assisted personal interview), usually administered in the respondent's home
- Interview language: German
- Target persons:
  - Anchor persons that ...
    - participated in PIAAC-L wave 2
    - were temporary non-participants from PIAAC-L wave 2
  - All household members in the anchor persons' household aged 18 and over
- Data collection period: March 1 to July 28, 2016
- Number of interviewers: 117
- Interview duration (on average, rounded):
  - Household protocol: 4 minutes
  - Household questionnaire: 9 minutes
  - Person questionnaire: 45 minutes
- Gross sample size (anchor persons): 3 510
  - Participants in wave 2: 3 263
  - Temporary non-participants in wave 2: 247
- Realized sample size:
  - Anchor persons: 2 967
  - Household members: 1 914
- Achieved retention rate: 84.5%
- Data: accessible for scientific purposes as scientific use files (ZA5989) from GESIS Data Archive / Research Data Centre PIAAC (FDZ PIAAC)
  - Latest version from 14.12.2017: GESIS – Leibniz Institute for the Social Sciences, German Socio-Economic Panel (SOEP) at DIW Berlin & LIfBi – Leibniz Institute for Educational Trajectories (2017): PIAAC-Longitudinal (PIAAC-L), Germany. GESIS Data Archive, Cologne. ZA5989 Data file Version 3.0.0, doi: [10.4232/1.12925](https://doi.org/10.4232/1.12925)

In chapter 2, some general remarks on features of wave 3 in the context of waves 1 and 2 are presented. Information on the survey instruments is provided in chapter 3, and chapter 4 summarizes fieldwork processes and results. Chapter 5 describes data management activities and data products.

## 2 Wave 3 in PIAAC-L

---

PIAAC-L pursues the achievement of a variety of goals. First of all, PIAAC-L offers the opportunity to collect additional information about the individuals that participated in PIAAC Germany and, by that, provides an enhanced basis for addressing research questions. The transfer of the cross-sectional German PIAAC sample into a longitudinal survey allows for the investigation of change and causality, e.g., with regard to skill maintenance, development, or loss. In addition, PIAAC-L was constructed to combine survey methodological knowledge, design features, and instruments from three large-scale surveys, PIAAC, the SOEP, and the NEPS. These surveys are conducted by leading institutions in the social sciences in Germany—GESIS, DIW Berlin, IIfBi—who have experience in the administration of panel and adult skills assessment surveys and contributed this expertise to the project. The decision to pool this expertise synergistically was the impulse for the implementation of PIAAC-L as a longitudinal project, including the application of a household concept (motivated by the SOEP approach) and a (re-)assessment of cognitive skills, using instruments from PIAAC as well as the NEPS.

### 2.1 Wave 3: The Final PIAAC-L Data Collection

PIAAC-L was not designed as a traditional panel survey. First, individuals or households were not initially sampled for the participation in a panel survey. Instead, respondents that participated in the cross-sectional PIAAC survey and agreed to be re-contacted for a follow-up survey were transferred into PIAAC-L and were the central response units in the panel. If such an anchor person became a final non-participant at some point in time, all of the household members also became dropouts, unlike the SOEP, where the eligible household members can continue to be interviewed even if, for example, the head of the household leaves the panel. Second, PIAAC-L has not been designed as a long-term panel, but is limited to a three-wave follow-up, with a slightly different design and focus in each wave.

In *wave 1* in 2014 every PIAAC respondent that agreed to be followed-up at the end of the PIAAC interview was considered as a target person for PIAAC-L. A household concept was applied, and by that, all household members of these target persons aged 18 or above were to be interviewed too. Core instruments from the SOEP, namely a household protocol, a household questionnaire, and a person questionnaire, were administered with only minor modifications in each participating anchor person household. For more information we refer to Zabal et al. (2016).

The central theme of *wave 2* in 2015 was a (re-)assessment of cognitive skills. Target units were anchor persons that participated in *wave 1* and—unlike in *wave 1*—only their partners, if these were living in the anchor person's household, regardless of their formal relationship status. Anchor persons were pre-allocated to one of eight assessment conditions. Two conditions implemented cognitive items from NEPS reading and mathematics only, two conditions administered tasks from the PIAAC literacy and numeracy domains only, the remaining four conditions were combinations of PIAAC and NEPS instruments. Partners took a NEPS assessment only, i.e. reading and mathematics. The cognitive assessment was preceded by a CAPI person questionnaire, including questions from a variety of source questionnaires (from PIAAC, SOEP, NEPS, and others). For more information see Zabal et al. (2017).

*Wave 3* in 2016 is the last component of PIAAC-L. Some design features resemble *wave 1*, such as the fact that target units were not only the anchor persons but again also household members 18 years or older. Once again, three instruments were implemented: a household protocol, a household questionnaire, and a person questionnaire. However, there were differences compared to *wave 1* and *wave 2*. On the one hand, a non-monotonic design was implemented. Some anchor persons did not participate in *wave 2* due to reasons that are temporary in nature, such as an absence from home during fieldwork in *wave 2* (e.g., stay abroad) but with the prospect of returning to their home by

2016. These *temporary non-participants* were fielded in 2016 (for more information see section 2.2 on the PIAAC-L sample composition). On the other hand, the wave 3 person questionnaire, which is again essentially a core SOEP instrument, has been extended by a set of additional, newly developed questions and two modules for the assessment of cognitive skills (one of them from the SOEP). Figure 2.1 summarizes the overall design of PIAAC-L across the three waves.

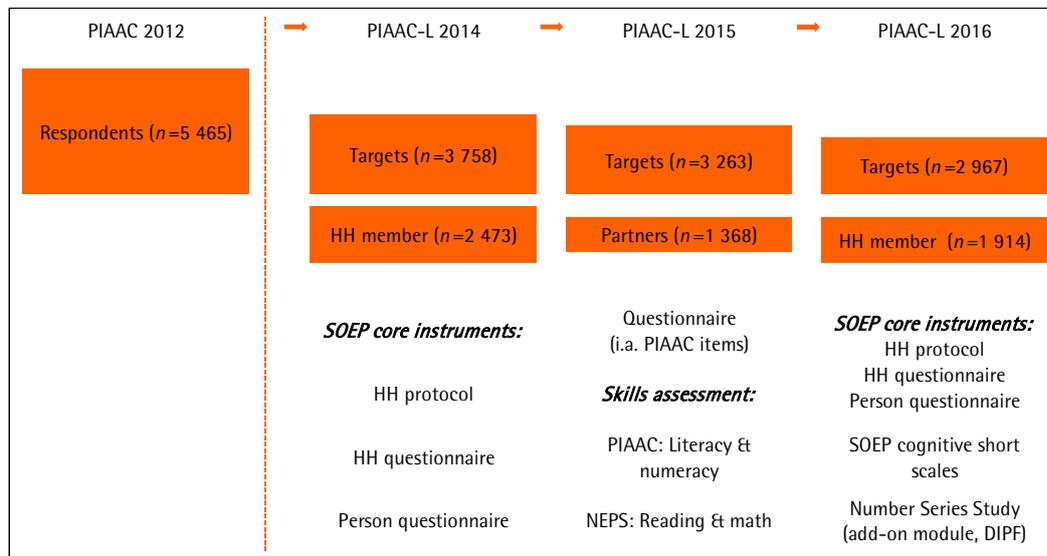


Figure 2.1. General overview PIAAC-L project

After completion of the PIAAC-L project, a part of the sample was transferred to the SOEP as a refreshment sample. These are anchor persons who explicitly gave their consent to being re-contacted within the scope of the SOEP. Some of these anchor persons have already been interviewed once in 2017 and will continue to be interviewed under the sole responsibility of DIW Berlin within the scope of the SOEP. For more information see Steinacker, Wolfert, and Thümmel (2017).

## 2.2 The PIAAC-L Sample in Wave 3

The starting point for the PIAAC-L anchor person sample is the initial sample from the German PIAAC survey 2011/12. A detailed description of the registry-based, two-stage stratified and clustered PIAAC sample design, the selection, the target population, and the sample size are given in Zabal et al. (2014). The technical report for PIAAC-L wave 1 (Zabal et al., 2016) summarizes relevant information on the initial sample for PIAAC-L.

The PIAAC-L wave 3 anchor person gross sample is composed of two groups of anchor persons from wave 2: respondents and temporary non-participants (see definition in the previous section). Altogether 3 263 anchor persons were respondents in wave 2 and thus eligible for wave 3. However, only 3 250 cases were fielded because at the end of the wave 2 interview or during the interview validation process 13 anchor persons reported their refusal to be further contacted for PIAAC-L. Following the non-monotonic design of wave 3, it was possible to re-contact anchor persons that were classified as temporary non-participants in wave 2. Overall 247 persons or around 50% of all non-participating anchor persons in wave 2 were classified as temporary non-participants. Table 2.1 summarizes the reasons for non-participation in wave 2 for this group of anchor persons.

Table 2.1. Reasons for Non-Interview in Wave 2 for Temporary Non-Participants

Reasons	<i>n</i>	%
Soft refusal	72	29.1
Other reason (unspecified) or unusual circumstances (e.g., temporary sickness, hospital stay)	63	25.5
Target person unavailable during field period (e.g., temporary absence from home due to work commitments in another city or a term abroad)	62	25.1
No contact during field period	42	17.0
Miscellaneous (e.g., permanent sickness)	8	3.2
Total	247	100.0

All 247 temporary non-participants were fielded in 2016. In summary, the gross anchor person sample for wave 3 encompassed 3 510 cases, of which 3 497 were fielded. About five years after their selection, 64% of the 5 465 PIAAC net sample cases were still panel members prior to the start of wave 3 fieldwork. Attrition between waves cannot be fully eliminated but in parts reduced through measures of panel maintenance. Anchor persons received a Season's Greetings card in December 2015. This measure had two main purposes: consolidation of the respondent's commitment to the survey and an address check. Invalid anchor person addresses were verified in an address tracking and updating process at the local registries prior to fieldwork.

At the outset of PIAAC, the sample was distributed across 277 communities within 320 sample points. Since then, some anchor persons moved to different communities and were successfully traced, resulting in a larger distribution of addresses across regions. At the beginning of the wave 3 fieldwork, 549 communities were included in the sample, which increased to 578 different communities at the end of fieldwork (Steinacker et al., 2017).

Members of participating anchor person households born 1998 or earlier were also eligible individuals in wave 3. According to the records in the household protocol, there were 3 164 additional eligible members in anchor persons' households. For the most part (61.8%), they were spouses and partners, followed by biological children (15.7%), biological parents (14.4%), biological siblings (4.8%), and persons with other types of relationships (3.3%). In technical terms, household members can be split up into three separate groups: (a) individuals that participated in at least one of the previous waves, (b) individuals that have been in the household in at least one of the previous waves but did not participate, e.g. children under age, and (c) individuals that were registered for the first time in wave 3. Whenever a member belonging to group (b) or (c) was eligible and willing to participate in the wave 3 interview, they were required to give their consent for linking their data basically with all the previously collected data of their household members during the course of PIAAC-L and PIAAC. This is in line with the terms of agreement that were established in wave 1 and is applicable for all respondents in PIAAC-L (see terms of agreement in Zabal et al., 2016, p. 9-10 as well as p. 12).

## 3 Instruments

---

The rationale in wave 3 was again to follow the SOEP design closely by interviewing all household members with a set of three instruments: the household protocol, the household questionnaire, and the person questionnaire. The content of the instruments was adapted to PIAAC-L desiderata and consisted only in parts of the SOEP core questions. At the end of the person questionnaire there were two additional modules for the assessment of cognitive skills, one of them taken from the SOEP.

The cognitive modules were designed as self-administered assessments, one of which was assigned only to selected anchor persons. The questionnaires were interviewer-administered as CAPI interviews. Ideally, in each household one household protocol, one household questionnaire and as many person questionnaires as eligible persons living in the household were to be completed. However, since the participation of the anchor person was the main objective, the completed case definition of wave 1 was applied. In addition to the household protocol and the household questionnaire, at minimum, the person questionnaire with the anchor person had to be administered (Zabal et al., 2016). To this effect, participation of other household members was not mandatory. Overall in 40 cases the completed case specification was not met, since at least one of the requested components was missing. Following the decision made in wave 1, 22 cases were accepted post-hoc as completed cases (only the household questionnaire was missing; in one case the household protocol was missing for technical reasons). In 18 cases the entire household had to be counted as a final loss and collected data had to be deleted because the anchor person interview was not obtained.

### 3.1 Contact Documentation and Household Protocol

In wave 1, contact documentation and household protocol were administered in the same instrument. For wave 3, data for these two components were collected in separate technical environments.

#### *Contact documentation*

Contact documentation in wave 3 was embedded in a tool that is accessible on the interviewer laptop for administrative-based interactions between interviewers and the survey organization's managers. Interviewers were requested to record the following information separately and in real-time for each contact/contact attempt which yielded a temporary (not final) disposition: type of contact (in person, by telephone etc.), date and time of contact, and result. A contact that led to a final disposition (e.g. interview or refusal), in contrast, had to be captured in the CAPI household protocol frame. If the anchor person household participated, the documentation of contacts/contact attempts was only pursued until the first interview, which in general was the administration of the household protocol. Subsequent contact activities undertaken to interview other household members were not recorded.

#### *Household protocol*

In wave 3, the household protocol had to be completed first (before all other instruments), both for content-related and technical reasons. It was mandatory to conduct the household protocol interview with the anchor person. The administration of the household protocol took on average four minutes.

The household protocol is an electronic form for recording information on the household composition. Data like name, year of birth, gender, relationship to anchor person, moving-in dates, information on presence and/or absence, and additional residences are captured for each member of the household, regardless of their eligibility. In year 3 of the panel, a large part of these data had already been collected in wave 1 or 2. Thus, existent information was pre-loaded into the CAPI questionnaire. The layout of the instrument enabled the interviewer to synchronize the pre-existing data with

information on the current situation in the anchor person's household. Changes to the household composition, updates or corrections to the data, or the registration of new household members were captured by the interviewer based on information provided by the anchor person.

Technically, other interviews in the household could only be unlocked after the completion of the household protocol. For this purpose, in the background, a sort of roster of eligible persons was created, based on the date of birth information. This roster information was electronically forwarded and used to activate a person questionnaire for each of the eligible persons. After finalizing the household protocol, no fixed sequence for conducting the household questionnaire and the person questionnaires was given. This was left to the interviewer's discretion, depending on the circumstances in the visited household.

If an anchor person or a member of his/her household refused participation in wave 3, the final disposition (according to a list of pre-specified reasons) was captured in the household protocol. Regardless of participation, interviewers recorded some paradata information, such as neighbourhood and dwelling characteristics, for each eligible anchor person household.

### 3.2 Household Questionnaire

The household questionnaire in wave 3 was based on the instrument implemented in the SOEP in 2015 (TNS Infratest Sozialforschung, 2015a); the SOEP source questionnaire was substantially shortened for PIAAC-L in content and length. While in wave 1 the average duration of the household questionnaire was 16 minutes, in wave 3 it was reduced to approximately 9 minutes.

Like in wave 1 per design it was intended that the household interview was to be conducted with the anchor person. However, in some cases the anchor person was not the household member that had the most comprehensive knowledge about details such as mortgage rates or utility bills. In 7% of the cases the interview was conducted with another household member, hereof two-thirds by the anchor person's partner and 30% by a parent.

The CAPI-based questionnaire was administered by the interviewer; show cards were provided to respondents for a small number of items. Content-wise the questionnaire was split up in four main sections:

- **Living situation, living conditions, and costs:** if the anchor person moved to the current house/apartment after being interviewed in 2014, information was collected on moving date, neighbourhood characteristics, and type of dwelling; characteristics of living space; ownership/tenancy, rent, mortgage, maintenance and additional costs
- **Household income and benefits, wealth:** household income and detailed sources; government aids and subsidies; savings; loans
- **Household possessions and customs:** item battery, including items on possessions (e.g. a landline phone), on current circumstances (e.g., whether there are financial savings for emergency situations), or on customary activities (e.g., going on a at least one-week-vacation every year); if an item was answered with "no" respondents were asked whether this was for financial reasons
- **Children:** information on year of birth and gender, school attendance, childcare and related costs was collected for up to six children born 1999 or later that were living in the household

In the annex, interested users can find a listing of questions that were kept in the PIAAC-L questionnaire compared to the SOEP source questionnaire. For a more thorough comparison of the PIAAC-L and SOEP instruments, users can obtain the documentation of the household questionnaire from the SOEP (TNS Infratest Sozialforschung, 2015a) and PIAAC-L 2016 (in German). The corresponding variables in the data set (ZA5989\_Household\_16) have English labels and are documented in the codebook.<sup>2</sup>

### 3.3 Person Questionnaire

The person questionnaire in wave 3 consisted of two parts, a core CAPI questionnaire and an assessment part. The core questionnaire was an instrument containing questions from different questionnaire sources (SOEP, PIAAC, the Adult Education Survey [AES]), and new developed questions. It was administered by the interviewer. For a number of questions respondents received show cards to select their responses. In a participating household, the anchor person and every household member born 1998 or earlier was eligible for a person interview.

After completion of the core part of the questionnaire, the assessment part started. It included two modules on the assessment of cognitive skills, and respondents solved the tasks primarily in a self-administration mode. Module 1 was administered to every person taking the person questionnaire, whereas module 2 was designed to be taken only by a subgroup of pre-selected anchor persons (for more information see sections 3.3.2 and 3.3.3). On average, the administration of the entire person questionnaire took approximately 45 minutes, which is about the same length as in wave 1.

#### 3.3.1 Core person questionnaire

The core person questionnaire covered questions from the following topics:

- **Background information, family, childhood, and cultural identity:** marital/civil status/partnership; born in Germany; cultural identity (identification with culture of origin, with German culture); citizenship; attendance of kindergarten or similar institution in childhood; parental information (birth, death, citizenship, education, occupation)
- **Biographical calendar:** retrospective information for the years 2014 and 2015 on participation in or occurrence of a number of events such as *in education, employed, retired, or homemaker*
- **Formal education (general and vocational education):** formal education (general and vocational); year of graduation; recognition of foreign qualification; current education; current vocational education; current further education
- **Adult education and training (vocational and general further training):** participation in training; number of training programs; topic, type, duration, employment status while participating in training; returns/benefits of training program; content of training (e.g., job-related); costs and coverage of costs; reasons for not attending a training program; information search on training programs
- **Work status, work situation and history, change of employer, skill mismatch:** current employment status and situation; unemployment; current occupation; current industry; company size; job changes; occupational status; working hours; side jobs; job resignation; skill mismatch (with respect to the respondents' assessment of how well their literacy and numeracy skills match what is required by their job; only for employed respondents)

<sup>2</sup> Household questionnaire (ZA5989\_fb\_Household\_16.pdf) and two codebook versions (ZA5989\_cod\_Household\_16.pdf and ZA5989\_cod\_Household\_16.xlsx) of PIAAC-L (wave 3) are accessible under <https://www.gesis.org/en/piaac/rdc/data/piaac-longitudinal> at the Research Data Centre PIAAC (FDZ PIAAC) or <https://dbk.gesis.org/DBKSearch/SDESC2.asp?no=5989&tab=3&db=E&tab=1> at the GESIS Data Archive.

- **Income and benefits:** current gross and net income; bonuses; benefits; income sources and income per source
- **Health, satisfaction, and attitudes:** current health; sleep; impairment; medical appointments; satisfaction with different areas; feelings; personality; attitudes towards life
- **Time use and leisure activities:** time use (on weekdays, Saturdays, Sundays)
- **Life events:** changes in family life events, such as *marriage, child was born, or father passed away* for 2014, 2015, and 2016 (indication of month of change in respective year)

Data which had already been obtained in one of the earlier waves from respondents with an interview in wave(s) 1 and/or 2 were no longer collected in wave 3. As a consequence some questions (e.g., on parental information) were only directed to respondents participating for the first time. Relevant filter instructions are documented in the person questionnaire documentation (in German). An overview of the variables of the corresponding data set (ZA5989\_Persons\_16) is provided in the codebook (in English).<sup>3</sup> Table A.1 in the annex shows the mapping of questions or groups of questions to their source questionnaire and indicates whether questions were adapted for PIAAC-L.

The person questionnaire in wave 3 also implemented questions on adult education and training, job changes, and skill mismatch that were contributed by members of the PIAAC Leibniz Network (PIAAC LN).<sup>4</sup> The module on *Adult Education and Training* aims at understanding the participation in adult education and individual barriers to participation. The PIAAC LN team developed a questionnaire module to measure further education and training in a meaningful and harmonized manner. This module consisted of approximately 40 questions, the majority of which were adapted versions of questions from the AES, the SOEP, and PIAAC, supplemented by some new questions.

A set of nine questions was developed to evaluate *job changes* and the year when the respondent started to work for an employer. The sequence started with a question related to the most recent employer and in an iterative process this information was collected for up to four additional previous employers.

Employed respondents were asked two newly developed questions measuring *skill mismatch*, i.e. how the respondents' self-assessed literacy and numeracy skills compare to what is required for their job. On a scale from 1 (low) to 11 (high), respondents had to assess their skills in literacy (first question) and numeracy (second question) in the context of their job requirements. Due to concerns that respondents could tend to overrate their skills if asked by the interviewer, respondents answered these questions in a self-administration mode. Thus, interviewers had to re-arrange their convertible interviewer laptop by swiveling the screen so that it could be used as a tablet. Respondents selected their answers by clicking on one of eleven radio buttons on the screen with a digital pen.

In addition, a question block on *cultural identity* with five questions was included in the person questionnaire.<sup>5</sup> Migrants were asked about their identification with their culture of origin and with the German culture. Non-migrants were asked about their identification with the German culture.

Prior to the implementation in the final person questionnaire, draft items were evaluated in a cognitive pretest (Otto et al., 2015). Qualitative interviews were conducted with 15 individuals who were selected based on a three-dimensional quota plan including gender (male/female), age group

<sup>3</sup> Person questionnaire (ZA5989\_fb\_Persons\_16.pdf) and two codebook versions (ZA5989\_cod\_Persons\_16.pdf and ZA5989\_cod\_Persons\_16.xlsx) can be accessed through the same website links provided in the section on the household questionnaire.

<sup>4</sup> The questions were developed by (a) Britta Gauly and Natascha Massing (GESIS) for *adult education and training*, (b) Franziska Hampf and Simon Wiederhold (ifo Institute) for *job changes*, and (c) Anja Perry (GESIS) for *skill mismatch*. In all cases the PIAAC-L team at GESIS cooperated with the PIAAC LN team and provided guidance and support. For more information on PIAAC LN see <https://www.gesis.org/en/piaac/projects/>

<sup>5</sup> The questions were developed by Débora B. Maehler (GESIS) in cooperation with the PIAAC-L team at GESIS.

(16-45/46+), and educational level (with/without higher education entrance qualification). The pretest covered the following topics:

- Four questions on cultural identity
- Two questions on attendance of kindergarten or similar institution in childhood
- Six questions related to the year when the respondent started to work for an employer

The results of the cognitive pretests were discussed with the item developers and all issues identified by the pretesting lab were addressed and reworked.

### 3.3.2 Basic cognitive skills (short scales)

As mentioned above, the assessment part of the interview included two modules assessing cognitive skills. In module 1, three short tests of cognitive ability were administered to all respondents:

- Animal Naming Task
- Multiple-Choice Vocabulary Intelligence Test
- Symbol-Digit Test

These tests are taken from the SOEP, where they were implemented in different combinations, for example in the core SOEP in 2006 (Heineck & Anger, 2008; Schupp, Herrmann, Jaensch, & Lang, 2008; TNS Infratest Sozialforschung, 2011) or the SOEP Innovation Sample (SOEP-IS) in 2014 (Bohlender & Glemser, 2016). Richter et al. (2017, p. 7-9) provide some details on the theoretical background and development of these tests as well as additional references for further information.

For the *Animal Naming Task*, a word fluency test, respondents had to mention as many different animals as possible within a 90-seconds time frame (Richter et al., 2017). For each word mentioned by the respondent, the interviewer recorded whether it was a correct animal name or species, a duplicate, or an unintelligible answer.

The *Multiple-Choice Vocabulary Intelligence Test* and the *Symbol-Digit Test* were both completed in a self-administration mode by the respondent on the laptop. The interviewer once again prepared the laptop by swiveling the screen in order to use it as a tablet. Answers were selected by clicking on the screen with a digital pen.

The *Multiple-Choice Vocabulary Intelligence Test*, as described by Richter et al. (2017), consists of 37 groups of five words each. Four of these words are fictive, and the respondent has to select the one word that truly exists. The degree of difficulty of the words increases from item to item. The test is not timed and takes about five minutes to complete (Richter et al., 2017).

For the *Symbol-Digit Test*, a test of perceptual speed, respondents were asked "to match as many numbers and symbols as possible within 90 seconds according to a given correspondence list which is visible to the respondents on a screen" (Heineck & Anger, 2008, p. 11). Examples for both, the *Multiple-Choice Vocabulary Intelligence Test* and the *Symbol-Digit Test*, are available in Bohlender and Glemser (2016).

### 3.3.3 The Number Series Study

In addition to the SOEP module measuring cognitive abilities, a number series test was embedded in the PIAAC-L wave 3 interview as an add-on module. The PIAAC LN partner institute DIPF (German Institute for International Educational Research) in Frankfurt is responsible for the contents of this test, the administration design, and the software tool. More information on the Number Series Study is provided in Engelhardt and Goldhammer (2018) and Steinacker et al. (2017).

The test was administered only to anchor persons that had taken a PIAAC-based assessment in literacy and/or numeracy in PIAAC-L wave 2 ( $n = 976$  of 1 091 pre-selected anchors). Of these 976 respondents, only 910 cases have number series data. Twenty-seven anchor persons (2.8%) refused to participate in the Number Series Study after completing the core person questionnaire. In 39 cases (4.0%), data from the number series test were either missing for technical reasons or were removed during data management (implausible data).

Prior to the number series test, anchor persons were informed about some details, such as that the DIPF is responsible for this module and the confirmation that the PIAAC-L group would share only anonymized data with them. Anchor persons who agreed to take this test worked in a self-administration mode with the laptop screen again swiveled into a tablet format. They received up to 15 different number series tasks in a 16-minutes time frame. For each task, a series of numbers was presented on the screen and the task was to complete that number series in a logical manner by entering one or two further numbers. A separate software tool was developed to implement this test and embedded within the survey organization's interview software. Entries were made with a digital pen; respondents selected their numerical answer from a digital number keyboard on the screen.

## 4 Fieldwork: Preparation, Operations, and Results

---

Data collection took place between March 1 and July 28, 2016. As in the previous waves, data collection and fieldwork processes were conducted by TNS Infratest (now: Kantar TNS), which was contracted by GESIS. Eight persons in the survey organization were in charge of the implementation and the management of PIAAC-L. TNS Infratest appointed one field director, who had the overall responsibility for the implementation and management of PIAAC-L. In addition there were seven supervisors responsible for controlling operations in the field and the direct monitoring of the interviewers' work and outcomes. The specifications for fieldwork processes and measures to be implemented by TNS Infratest were specified by GESIS, which was also responsible for monitoring and controlling TNS Infratest's work.

The present section is subdivided into three main parts on preparatory work prior to fieldwork, summary of field operations, and fieldwork results. Further information on fieldwork can also be found in the TNS Infratest fieldwork report for wave 3 (Steinacker et al., 2017).

### 4.1 Preparatory Work

Preparatory work prior to the fieldwork start include the selection, remuneration and training of interviewers, the development of supporting material for interviewers, and the selection of adequate measures for addressing respondents.

#### *Interviewer selection, remuneration, and training*

PIAAC-L as well as PIAAC are highly demanding surveys that include the administration of assessment components, partly embedded software tools unfamiliar to interviewers, and have a complex design (e.g., panel with anchor person concept and varying groups of target respondents per wave). The *interviewers selected* for PIAAC-L for wave 3 were therefore ideally expected to have experience in conducting PIAAC-L in wave(s) 1 and/or 2 and if possible also have knowledge of PIAAC. For wave 3, TNS Infratest appointed 117 interviewers: 98 of them with PIAAC/PIAAC-L experience, 15 only with PIAAC-L experience,<sup>6</sup> and four interviewers who were new to the project and had neither experience with PIAAC nor PIAAC-L (Steinacker et al., 2017). However, these four interviewers had expertise with the administration of panel surveys, e.g., the SOEP. Whenever possible, interviewers were allocated to the anchor persons they had interviewed in the previous wave(s). In some instances, however, an interviewer change was inevitable, e.g., when the anchor person moved to another municipality that was no longer within the working radius of the former interviewer.

The PIAAC-L wave 3 interviewer pool included more male interviewers than female interviewers (54% vs. 46%). Around 71% of the interviewers were 60 years or older, 8% were below 50 years of age. The majority of the interviewers (61%) had a lower secondary degree (Mittlere Reife) or below (Hauptschulabschluss). About 89% of the interviewers had been working for TNS Infratest for five or more years, 29 interviewers (25%) for 15 years or more.

The *interviewer remuneration* scheme in wave 3 was in line with the previous waves of PIAAC-L. In comparison to other national surveys of a comparable scope, the payment in PIAAC-L was somewhat more attractive. The basic salary was based on a completed case (including a household protocol interview, a household interview, and an anchor person interview). Every additional person interview with an anchor person's household member and the administration of the Number Series Study were paid separately. A bonus was provided if all eligible household members were successfully interviewed.

---

<sup>6</sup> One of these, however, worked as an interviewer in the PIAAC Field Test in 2010.

For exceptional cases, e.g., if a fairly long drive to a sample point was involved, an extra amount was paid. Allowable expenses, such as costs for phone calls or travel expenses (e.g., gas), were refunded according to common TNS Infratest practice.

In each PIAAC-L wave, special project-related *interviewer trainings* were conducted. These supplement the basic interviewer training provided by TNS Infratest. In wave 3, a face-to-face interviewer training was considered necessary to address wave 3 design features, the specifics of the person questionnaire, and the cognitive test components. There were four separate training sessions (3.5 to 4 hours in length) at different locations, each with around 30 participating interviewers. The trainings were led by TNS Infratest with the support of the PIAAC-L group (GESIS).<sup>7</sup>

Three interviewers could not personally attend any of these meetings. Thus, they were trained by TNS Infratest supervisors through WebEx (online videoconferencing application). In detail, the following topics were covered in each training session: project design 2016, including instruments, material, and specifics; overview person questionnaire with special focus on new questions; administration of cognitive tests; and miscellaneous, such as contact documentation, quality control, and remuneration.

#### *Supporting material for interviewers*

In addition to the training sessions, interviewers were prepared for the upcoming fieldwork with supporting material. Like in previous waves, they received an interviewer manual (mainly developed by TNS Infratest) that summarized all relevant features of the third wave, such as:

- General information
- Project design
- Structure and design of the wave 3 survey
- CAPI instruments and administration of cognitive tests
- Practice interviews
- Case documentation and disposition codes
- Material
- Technical details about laptops
- A register listing and explaining crucial terms used in the questionnaires (adopted from the SOEP)
- Information on the migration of the sample after termination of PIAAC-L

Prior to the start of the fieldwork period interviewers were equipped with the following material:

- Convertible laptop with a digital pen for the administration of the cognitive tests
- Contact protocols for each anchor person household, list of addresses to be worked
- Contact cards (project specific)
- Advance letters (two versions)
- Confidentiality/data privacy information sheets
- Information sheet for new respondents on obtaining consent for linkage of data
- Information sheet for anchor persons on the planned migration of the sample after the end of PIAAC-L
- Interviewer manual
- Show card booklet

To get acquainted with the survey instruments and the wave 3 processes, interviewers were encouraged to conduct some test interviews prior to their first contacts or contact attempts.

---

<sup>7</sup> One of the interviewer trainings was attended by a researcher from the DIPF to introduce their add-on module.

### *Measures for addressing respondents*

Given that this is the third wave of the PIAAC-L survey, a rapport between interviewers and their anchor person households had already been established. These households were familiar with the PIAAC-L project and the reality of being contacted once a year for additional interviews. Thus, it was not as crucial to implement a broad set of measures to gain cooperation compared to earlier waves of PIAAC-L or at the outset of PIAAC.

Due to the fact that some anchor persons temporarily did not participate in wave 2 in 2015, two separate *advance letters* were prepared. One letter version was sent to anchor persons that had participated in 2015 and the other version to anchor persons who temporarily did not participate in 2015 (examples of these letters are provided in Steinacker et al., 2017). Advance letters were sent ahead of the fieldwork start to anchor persons by mail and were accompanied by a confidentiality/data privacy information sheet. By mailing the letters prior to the fieldwork start, it was also possible to check whether the anchor persons' addresses were still up-to-date. If letters were returned as "undeliverable," TNS Infratest started an address search with the local registry in order to obtain the new address of the respondent. In contrast to previous waves, no advance letters were sent to cases that needed to be re-worked in the re-issue phase, especially because the number of cases to be considered was small anyhow (see more below in section 4.2).

Providing conditional *incentives* to participants has a tradition in PIAAC and PIAAC-L. The burden that is imposed on respondents and their households due to long interviews with assessment components and the panel character of the study demand an adequate incentive strategy. Each wave has a somewhat different focus and design, resulting in different incentives per wave. Compared to wave 2, the incentive in wave 3 was reduced due to a lower average interview time. In comparison to wave 1, however, the incentive in wave 3 was higher because the person interview also included assessment components. In wave 3 the following incentive scheme was used:

- For a completed case: 30 euros (5 euros more than in 2014)
- For each additional person interview with an eligible household member: 20 euros (10 euros more than in 2014)

There was no additional incentive for anchor persons who took part in the Number Series Study.

## **4.2 Field Operations**

In the present section, the different fieldwork phases are presented, followed by specifics about fieldwork procedures. Information on the quality control activities completes this section.

### *Fieldwork phases*

Fieldwork was split up in a main working and a re-issue phase. In some cases, however, no clear cut was made between the two phases, especially if appointments were already scheduled for a date that technically fell into the re-issue phase. The two phases refer to the first interview to be conducted in the target household, namely the household protocol interview with the anchor persons. No separate working phases were set up for follow-up attempts for further interviews, e.g., with other household members. Supervisors from TNS Infratest monitored interviewers and made sure that a final result (either an interview or a code for non-participation) was captured by the interviewers for each eligible household member.

The *main working phase* took place between March 1 and July 27, 2016. A total of 3 497 anchor person addresses were released at the beginning of fieldwork. The majority of these cases were released simultaneously. For a few cases, however, interviewers requested to start with a small delay. Interviews with 2 907 anchor persons (83%) were obtained in this main working phase.<sup>8</sup> Reasons for non-participation (590 cases) were:

- Refusals (~ 6.5%)
- Non-contacts (~ 2.5%)
- Address-related reasons (~ 2.4%)
- Not available during fieldwork period (~ 1.6%)
- Ineligible, e.g. moved abroad or deceased (~ 0.4%)
- Language- or health-related reasons (~ 0.4%)
- Other reasons (~ 3.0%)

The reasons for non-participation in the main phase were scrutinized by TNS Infratest. For example, neither ineligibles nor individuals with a hard refusal (e.g., due to data privacy concerns) could be re-contacted. In total 167 cases were released for re-working attempts in the re-issue phase.<sup>9</sup> Anchor persons who temporarily dropped out in 2015 and did not provide an interview during the main working phase were not re-issued.

The re-issue phase was scheduled for the period of June 30 to July 28, 2016. With regard to non-participation due to an incorrect address, only cases for which a new address was obtained in the course of June 2016 were formally re-issued. Moreover, new address information obtained throughout the main working phase—either from interviewers directly or through an update from a local registry—was passed to the field on a continuous basis (i.e. independent of working phases). Sixty (36%) of the 167 re-worked cases provided an interview in the re-issue phase.

#### *Specifics about fieldwork procedures*

As mentioned in the previous subsection, problems in contacting respondents due to inaccurate addresses is smaller than in cross-sectional surveys but still an omnipresent issue in panels. Thus, *address searches* at different stages prior to and during data collection are necessary. TNS Infratest reported different points in time and sources for obtaining information that addresses were no longer up-to-date in wave 3 (Steinacker et al., 2017):

- Information from wave 2, either from interviewers during data collection (when they became aware of a planned move) or from respondents in the course of the quality control process
- End of 2015/begin of 2016, as a result of the Season's Greetings card sent by postal mail to anchor persons' households (panel maintenance activity)
- Prior to fieldwork in wave 3, as a result of the advance letters sent by postal mail
- During fieldwork in wave 3, from interviewers

<sup>8</sup> As mentioned in section 3, disposition codes for an interview or a reason for nonresponse were recorded by the interviewer for each contact or contact attempt up to the initial interview with the anchor person (household protocol). Technically, the reported number of successful interviews refers to these completed initial interviews with the anchor person. However, in a retrospective view, a successful interview also symbolizes a completed case according to the PIAAC-L definition.

<sup>9</sup> Of these 167 cases, 32% were non-contacts, 31% had address-related issues, 21% were (soft) refusals, 11% were not available during fieldwork period, and 5% did not participate due to other reasons.

Depending on the source, new addresses were usually obtained through updated information from the post office, search requests at local registries, or independent research by interviewers. For 59 cases a search request was started at a local registry, resulting in 52 new addresses that could be distributed to interviewers. In 87% of these cases an interview was conducted.

Best practices in *contacting*, such as the distribution of contact attempts across different days and times, as described for wave 1 in Zabal et al. (2016), were also a guideline for interviewers in wave 3. The protocol was relaxed for wave 2 and continued analogously in wave 3 following the specifications of wave 2. At this stage in the panel, survey interviewers had already established a relationship with anchor persons and their households. They were therefore familiar with individual situations in those households and could predict adequate times for contacting. In addition, they usually had up-to-date contact information, including phone numbers. Thus, for wave 3 initial contact attempts were permitted via phone. In cases when contact attempts by phone were not successful, however, at least four in-person attempts were required. In the re-issue phase, potentially difficult cases had to be contacted in-person. Contact history information and temporary (non-final) results of contacts and contact attempts were recorded electronically (see section 3.1).

Like in the past waves and in PIAAC, TNS Infratest offered a toll-free *hotline* for respondents during regular office hours. Respondents could directly address their questions or concerns to the PIAAC-L field director at TNS Infratest. Altogether only 48 incoming calls were registered, a number that is considerably lower than in the past. The technical hotline for interviewers, in contrast, was contacted 170 times; 68 interviewers consulted the technical hotline (Steinacker et al., 2017).

#### *Quality control*

Interview validation and control of key information were essential elements of the quality control process implemented in PIAAC and PIAAC-L. In essence, well-established quality control procedures in validating interviews with anchor persons were continued in wave 3, such as (cf. Zabal et al., 2016): (a) 100% validation of interviews by mailing a validation questionnaire to anchor persons, (b) consistency checks for year of birth, gender, and first name with information on the returned validation questionnaire, interview data and preload information,<sup>10</sup> and (c) checks, whether interview duration seemed plausible.<sup>11</sup>

For the postal-based control process, a brief validation questionnaire and a letter with instructions were sent to anchor persons shortly after they completed their interview. Anchor persons were asked to answer the questions and return the questionnaire to TNS Infratest using the enclosed self-addressed envelope. The return rate was 60%. The questionnaire covered questions on:

- whether the anchor person had been recently interviewed by an interviewer for PIAAC-L
- whether the anchor person had been asked by the interviewer to work independently on some exercises on the laptop
- whether the interviewer helped the anchor person in solving the exercises
- duration of the interview
- number of household members born 1998 or earlier
- number of those reported household members who gave an interview
- year of birth and gender of anchor person

The two questions on the number of household members were embedded to validate whether the registered number of additional person interviews per household was congruent with the information

---

<sup>10</sup> In 2014 and in PIAAC the checks were based on information from the sampling frame (i.e., registry data).

<sup>11</sup> Interviews with a duration of 30 minutes or below (including the administration of the number series test) and of 22 minutes or below (without the number series test) were further investigated (Steinacker et al., 2017).

provided by the anchor person. In fact, a number of cases were identified in which anchor persons seemed to have misinterpreted these questions. In a number of cases the difference was exactly one and further consultation with the anchor persons revealed that they had included themselves in the total number. Thus, later on in the validation process only cases in which the difference was equal to or greater than two were further investigated. After that adjustment the number of "conspicuous" cases declined.

In addition to that, the criteria for flagging cases for further checks followed the spirit of validation checks in 2014 (Zabal et al., 2016) and 2015 (Zabal et al., 2017), such as discrepancies in year of birth or gender, short interviews, but also indications that no cognitive tests were administered or that the interviewer helped the respondent solve the cognitive exercises. All information and documentation on cases with conspicuous patterns was systematically captured, and TNS Infratest followed a strict protocol to scrutinize these cases, including clarification by directly contacting the anchor person by phone. Results were shared with GESIS, and for a number of individual cases a consultation between TNS Infratest and GESIS took place. Initially 282 cases were flagged and further investigated. The number was reduced to 149 after the exclusion of the above-mentioned cases with a deviation of one for the number of participating household members. For around two-thirds of these cases the observed discrepancies were resolved in a phone call with the anchor person. The remaining cases were intensively examined using all available data and interviewer feedback. In this context, TNS Infratest also checked whether individual interviewers with a cluster of issues could be identified. This was not the case, however, and eventually all cases were cleared.

Due to the fact, that anchor persons are the central response unit, person interviews with other household members were not systematically included in the validation process. However, the potential for falsification cannot be ignored for these interviews. On the one hand, the questions on the number of eligible and participating household members in the anchor person's validation questionnaire were one source of information of that verification. On the other hand, one new and additional control feature was implemented in wave 3: For each interviewer, a household with at least one non-anchor person interview was randomly selected, contacted and one additional interview with a household member validated. The selection of individuals within a household ensured that both partners and other household members had a chance to be selected.<sup>12</sup> TNS Infratest staff called the person with whom the interview had been conducted and checked whether the interview had actually taken place using a validation questionnaire (Steinacker et al., 2017) verifying first name, gender, age, date, time and duration of interview as well as some background information on the household. Successful calls were made for 110 interviewers and the administration of the interview was confirmed. Thus, this additional verification was not achieved for only very few interviewers. This is mainly due to the fact that no phone numbers were available for certain households or the household could not be reached during the validation period.

---

<sup>12</sup> Partners were contacted in 56% of the cases; the remaining 44% were other household members.

### 4.3 Fieldwork Results

A completed case in PIAAC-L included at least a household protocol, a household interview, and, at minimum, a person interview with the anchor. In wave 3, data are available for 2 967 completed cases at the anchor person level. In total the following numbers of interviews were completed in the corresponding anchor person households:

- 2 966 household protocols<sup>13</sup>
- 2 946 household interviews<sup>14</sup>
- 4 881 person interviews

At the beginning of wave 3, the gross sample comprised 3 510 anchor persons (and their households). The majority, 3 263 anchor persons, had also participated in wave 2; another 247 persons were temporary non-participants in wave 2, but could be re-contacted in wave 3 (see section 2.2). The simplest form of retention rate calculation in wave 3 is based on the division of completed cases by the gross sample for wave 3, thus 2 967/3 510, resulting in 84.5%. For anchors who participated in all three PIAAC-L waves the retention rate is 2 900/3 263 or 88.9%, compared to temporary non-participants, for whom the retention rate only reaches 27.1% (= 67/247).

#### *Outcomes of anchor persons*

When looking at the development of the anchor person sample from its starting point at the end of PIAAC 2012 until the end of PIAAC-L wave 3 (see Table 4.1), it can be seen that 54.3% of the 5 465 respondents of the PIAAC 2012 net sample are still panelists at the end of PIAAC-L. The largest attrition occurred during the transition from PIAAC 2012 to PIAAC-L wave 1 with a loss of 1 707 anchor persons (31%). Hereof, 240 participants of the PIAAC 2012 survey (4%) were not eligible target persons at the outset of PIAAC-L and thus, only 5 225 were included in the initial PIAAC-L wave 1 gross sample. In relation to this number of cases, 56.8% of the PIAAC respondents were still members of the PIAAC-L panel at the end of wave 3. The retention rate as observed for wave 1 and wave 2 increased from 71.9% to 86.8%. Due to the non-monotonic design, the retention rate of 84.5% in wave 3 is not directly comparable to these previous rates. However, if the rate of 88.9% (as reported in the previous paragraph) is used for comparison, it can be concluded that the retention rate reached a fairly stable benchmark.

---

<sup>13</sup> For one case no household protocol could be administered due to technical problems.

<sup>14</sup> In 21 cases no household interview was administered. However, like in wave 1, the completed case definition was relaxed if at the least the anchor person interview in the household was obtained (Zabal et al., 2016).

Table 4.1. Development of Anchor Person Sample From PIAAC 2012 to PIAAC-L Wave 3

<b>PIAAC 2012 net sample</b>	<b>5 465</b>
Anchor persons with consent for a re-contact in PIAAC-L	5 225
<i>Percentage (in relation to PIAAC 2012 net sample)</i>	<i>95.6</i>
<b>PIAAC-L 2014 initial gross sample</b>	<b>5 225</b>
Anchor person interviews in 2014	3 758
<i>Percentage (in relation to PIAAC-L 2014 initial gross sample)</i>	<i>71.9</i>
<i>Percentage (in relation to PIAAC 2012 net sample)</i>	<i>68.8</i>
<b>PIAAC-L 2015 gross sample</b>	<b>3 758</b>
Anchor person interviews in 2015	3 263
<i>Percentage (in relation to PIAAC-L 2015 gross sample)</i>	<i>86.8</i>
<i>Percentage (in relation to PIAAC-L 2014 initial gross sample)</i>	<i>62.4</i>
<i>Percentage (in relation to PIAAC 2012 net sample)</i>	<i>59.7</i>
<b>PIAAC-L 2016 gross sample</b>	<b>3 510</b>
Anchor person interviews in 2016	2 967
<i>Percentage (in relation to PIAAC-L 2016 gross sample)</i>	<i>84.5</i>
<i>Percentage (in relation to PIAAC-L 2014 initial gross sample)</i>	<i>56.8</i>
<i>Percentage (in relation to PIAAC 2012 net sample)</i>	<i>54.3</i>

The final outcome for nonresponse of the anchor person was recorded by the interviewer when s/he could not obtain consent from the anchor to administer a household protocol. To determine the final nonresponse disposition code, the entire body of information from the contact phase was processed. In consequence, the most informative and thus not necessarily the last information was used. For example, if an appointment for an interview was made with an anchor during the main phase, but was later cancelled by the anchor and the interviewer could not establish any further contact during the re-issue phase, the final outcome is nonresponse due to refusal rather than non-contact. The final disposition codes for the anchor persons, separately for wave 2 participants, temporary non-participants, and the entire sample, are shown in Table 4.2.

Even in the third wave of the panel, refusals were the most common reason for nonresponse. This is likewise valid for wave 2 participants as well as temporary non-participants. The second most reported reason for non-participation for both groups is related to unspecified other reasons/unusual circumstances. Due to the fact that only 27% of the temporary non-participants completed an interview, the proportions for the two above-mentioned reasons for nonresponse were higher in this group compared to the proportions for wave 2 respondents. A non-contact as reason for non-participation appeared to be more common among temporary non-participants and was of less relevance among wave 2 respondents. The latter is not surprising as interviewers were more familiar with the circumstances and patterns of presence in the participating households. The remaining reasons for nonresponse did not play a significant role. Furthermore, address-related issues were no longer of substantial importance, because the address list was well maintained in this panel with fairly short intervals between waves.

Table 4.2. Final Disposition Codes for Anchor Persons

Final disposition code	W2 R <i>n</i>	%	W2 TNP <i>n</i>	%	Total <i>n</i>	%
AP unavailable during fieldwork period	30	0.9	11	4.5	41	1.2
AP person permanently ill or incapable	11	0.3	2	0.8	13	0.4
Language problem, German proficiency insufficient	1	<0.1	1	0.4	2	0.1
Refusal	164	5.0	83	33.6	247	7.0
AP moved outside of country	5	0.2	3	1.2	8	0.2
AP moved, new address unknown	31	1.0	12	4.9	43	1.2
AP moved, new address known	6	0.2	3	1.2	9	0.3
AP unknown at given address	0	0.0	1	0.4	1	<0.1
Invalid address	2	0.1	2	0.8	4	0.1
Not a residential address, dwelling not inhabited	0	0.0	1	0.4	1	<0.1
Death	8	0.2	0	0.0	8	0.2
Other reasons or unusual circumstances	69	2.1	31	12.6	100	2.8
Non-contact	23	0.7	29	11.7	52	1.5
CC: HH protocol and interview, AP interview	2 878	88.2	67	27.1	2 945	83.9
CC: HH and AP interview, no HH protocol	1	<0.1	0	0.0	1	<0.1
CC: HH protocol and AP interview, no HH interview	21	0.6	0	0.0	21	0.6
Interview not usable	13	0.4	1	0.4	14	0.4
Total	3 263	100.0	247	100.0	3 510	100.0

Notes: W2 R = Wave 2 respondents; W2 TNP = Temporary non-participants in wave 2, AP = Anchor person; HH = Household; CC = Completed case

### Outcomes of eligible household members

In wave 3 a household concept was re-applied, and the goal was to obtain as many person interviews in the anchor person household as possible with household members born 1998 or earlier. In 20% of the cases anchors lived alone, while 36% lived in a two-person, 20% in a three-person, 18% in a four-person, and 6% in a multi-person household. Interviewers recorded a total of 3 164 other eligible persons in the anchor person households. Thus, on average, there were 2.07 eligible persons (including the anchor) per household. Altogether, 1 914 interviews with other household members were obtained. This is a completion rate of 60.5%, which is slightly higher than the corresponding rate in 2014 (= 58.7%; see Zabal et al., 2016). Table 4.3 illustrates disposition codes for other eligible household members.

Table 4.3. Final Disposition Codes for Eligible Household Members

Final disposition code	<i>n</i>	%
Unavailable during fieldwork period	64	2.0
Permanently ill or incapable	32	1.0
Language problem, German proficiency insufficient	21	0.7
Refusal	876	27.7
Other reasons or unusual circumstances	98	3.1
Non-contact	12	0.4
Interview not usable	147	4.6
Interview	1 914	60.5
Total	3 164	100.0

By far the most common reason for non-participation among other household members was a refusal (27.7%). Technical issues, which resulted in incomplete and non-usable interviews, as well as unspecified other reasons or unusual circumstances and absence during the fieldwork period added up to approximately another 10%. The commitment to participating in the third wave of PIAAC-L varied across groups of household members. Household members that are either partners or biological relatives (such as children or parents) of the anchor and who already participated in 2014 showed the highest commitment, with participation rates of 75% or more. Children or other household members who were too young to be eligible in 2014 but became eligible in 2016 had relatively high participation rates, ranging from 50% to less than 75%. Hard-to-convince household members were mainly those who already avoided participation in any of the previous waves. Participation rates for this group of people are below 25%, with the lowest rate of only 9% for partners without participation in both, 2014 and 2015.

#### *Interview duration*

Average interview duration times (in minutes) is 4 minutes for the household protocol, 9 minutes for the household interview, and 45 minutes for the person questionnaire. In a single anchor person household the overall interview length, on average, added up to approximately one hour. Given the fact that the person interview was the most time-consuming one, the burden in a household increased almost linearly with the number of participating persons, e.g., a household with two respondents had on average a total interviewing time of approximately 1.75 hours and so on. In general, anchor person interviews were longer than interviews with other household members (49 minutes vs. 40 minutes, on average). Anchor persons who were administered tasks from the Number Series Study spent approximately 56 minutes (on average) on the person interview.

## 5 Data Processing and Data Products

---

To a large extent, data processing activities in wave 3 resembled those in wave 1. After data collection, several stages of data cleaning and processing were carried out by the survey organization and by the PIAAC-L project team at DIW Berlin and GESIS. In this phase, GESIS was responsible for checking the data cleaning activities of the survey organization and updating the registry data set. DIW Berlin was in charge of completing all other data sets, with the exception of the Number Series data set. Coding of responses in open-text format was undertaken by the LIfBi. Later in the process, weighting, the computation of derived variables and the production of questionnaire documentation and codebooks was carried out at DIW Berlin. GESIS took over the thorough examination of all data sets and documentation (supported by the LIfBi), implemented the data protection measures in the data and was responsible for the data release. Cleaning and preparation of data (like labeling or computation of derived variables) resulting from the Number Series Study was carried out by the DIPF.

### 5.1 Data Cleaning, Processing, and Coding

Wave 3 followed a SOEP design and used core questions from the SOEP instruments, extended by newly developed questions. In order to ensure comparability with data released in the SOEP, PIAAC-L data cleaning and processing relied on well-established SOEP standards at the survey organization as well as the DIW Berlin. However, special care was needed in data processing since the PIAAC-L instruments only covered a limited number of questions used in the SOEP at the same time and used partly different filters. In addition, PIAAC-L specifics, such as the missing scheme, had to be applied to the data. Therefore, existing software codes at the survey organization and at the DIW Berlin were adapted or extended where necessary, and by that, were prone to error. Several iterations of data checking were required. All errors detected were subsequently resolved.

Another source of error stemmed from longitudinal inconsistencies. First, the composition of each household had to be verified across waves. Given the information on anchor persons and their household members from the previous waves, the household composition recorded by interviewers in wave 3 was synchronized with that information. Actual records might have indicated a change of the household composition due to the fact that persons left the household or new ones came into the household. New household members received serial numbers that were not previously assigned to any other person in the household. Serial numbers of persons who left the household were blocked for further use. This synchronization was part of the survey organization's work and reviewed by the project team at GESIS.

Second, the measurement of identical information across waves may have resulted in inconsistencies. The slightly different design and focus in each PIAAC-L wave resulted partially in the repeated collection of certain information (e.g., gender, year of birth, country of birth, or educational attainment). Although variables like year or country of birth should contain invariant information, PIAAC-L, like other panels, had to deal with the fact that there were in some few cases inconsistent data across waves. With regard to educational attainment inconsistencies may also be induced because questions were worded differently across waves. There were no data correction processes, however, in particular because it is usually not possible to determine the "true" answer. Moreover, PIAAC-L data reflect the results as collected in the field. It is our policy to leave it to the discretion of each PIAAC-L data user to decide how to deal with these inconsistencies.

Like in wave 1, coding of open responses on occupation, industry, and some questions on educational attainment was carried out by the coding department at the Research Data Center at the LIfBi. The procedure that was implemented in wave 1 was replicated in wave 3. The open responses on

occupation were directly coded into the "Classification of Occupations 2010" (KldB-2010). Codes for the "International Classification of Occupations 2008" (ISCO-08) were subsequently obtained using a crosswalk. In cases where the open-text information could not be converted directly via crosswalk, specially trained coders encoded natively in ISCO-08. Indices, such as ISEI-08, SIOPS-08 and MPS were then derived from ISCO-08. Information on industry was directly coded into the "International Standard Industrial Classification of all Economic Activities" (ISIC Rev. 4) and the "Statistical Classification of Economic Activities in the European Community" (Nomenclature statistiques des activités économiques dans la Communauté européenne – NACE Rev. 2) was subsequently derived. For more details on the coding process in wave 1 see Zabal et al. (2016).

## 5.2 Measurement of Basic Cognitive Skills in Wave 3

The third wave of data collection for PIAAC-L included two modules assessing cognitive skills which are described in more detail in section 3: the short scales measuring basic cognitive skills and the Number Series Study.

Data set ZA5989\_Cognit\_16 includes data resulting from the three short tests of cognitive ability that were taken from the SOEP. Data management activities for the Cognit data set included handling incomplete or invalid data, assigning missing values, and re-coding variables wherever needed. Whenever a case lacked valid data for all three tests, due to either technical problems or interviewer intervention, it was excluded from the data set. If data were missing for only one of the tests (e.g., due to partial refusal) the missing code -7 "Implausible value or not determinable" was assigned. See variables *flag\_validdata\_16*, *flag\_reason\_16* and *flag\_symbol\_16* in data set ZA5989\_Persons\_16 for more information. In addition, DIW Berlin computed some derived variables summarizing information.

Data from the Number Series Study is stored in data set ZA5989\_NumberSeries\_16. The data were checked and processed by colleagues from the DIPF. Data is available for 910 PIAAC-L anchor persons (see also variable *number\_series\_16* in data set ZA5989\_Persons\_16). More information on the Number Series Study is provided in Engelhardt and Goldhammer (2018) and Steinacker et al. (2017).

While PIAAC 2012 assessment data for literacy, numeracy, and problem solving in technology-rich environments and background information from PIAAC 2012 and PIAAC-L waves 1 (and 2) were included in IRT models to produce a new set of plausible values for PIAAC-L wave 1 (and 2), scaling of plausible values was no longer an option for wave 3. This is due to limitations in the estimation of IRT models. In order to enable users to run their analyses, the LfBi developed a tool based on R code (R package "PVPIAACL: Plausible Values estimation with the PIAAC-L data", available on <https://github.com/jcgaasch/PVPIAACL>) that allows users to estimate plausible values for their specific research purposes, i.e. by including exactly the set of background variables relevant for the users' individual analyses. For more information see Carstensen, Gaasch, and Rothaug (2018).

### 5.3 Weighting

In general, weighting procedures, as implemented in wave 1 and 2, were continued in wave 3. Weighting factors, namely nonresponse weights (variable *bleib\_16*) and poststratification weights (variable *hrf\_16*), were computed for anchor persons only and are stored in data set ZA5989\_Weights\_16. At first, probabilities were modeled in two steps, differentiating between reasons for non-contact vs. non-participation due to e.g. refusals, unavailability during field period, or unusual circumstances. A special feature of wave 3 was the treatment of temporary non-participants in wave 2, as these cases had missing values for variables from wave 2 which were used as predictors in the models. Multiplication of probabilities from each model yielded staying probabilities. The reciprocal value of these probabilities resulted in nonresponse weights (= inverse staying probability). In the subsequent calibration step, some central variables were adjusted to external benchmark data from the German Microcensus 2016. For a detailed description of the weighting procedure in wave 3, the treatment of temporary non-participants, and the combination of weighting factors from different waves for longitudinal analyses see Burkhardt, Silbermann, and Bartsch (2018).

Variance estimation to account for the complex survey design using replicate weights like in PIAAC 2012 was discontinued in PIAAC-L. Thus, it is suggested to estimate standard errors through Taylor Series Linearization by using variables on sampling and stratification from the PIAAC 2012 Scientific Use File (GESIS Data Archive, Data File ZA5845, current version 2.2.0, doi: [10.4232/1.12660](https://doi.org/10.4232/1.12660)). An exemplary code in Stata is provided in Burkhardt et al. (2018) and the document "*Notes to the User*" (access information see below).

### 5.4 Data Dissemination

Final steps in the data management process included the implementation of data confidentiality edits and the finalization of user-friendly data sets with accompanying documentation. An overall data confidentiality strategy for PIAAC-L was elaborated in wave 1 (Zabal et al., 2016) and extended in wave 2 (Zabal et al., 2017) and in wave 3. All measures were approved by data protection officers and principle investigators at each partner institute. For a large number of variables, measures defined for wave 1 and wave 2 could simply be re-applied in wave 3. Variables that were new had to be screened case-wise and for each one it was evaluated whether it bore a disclosure risk. Open-text responses and a few other variables were suppressed or coarsened as a consequence of this evaluation process. GESIS implemented the data confidentiality edits which were checked by the other PIAAC-L partners.

During the course of PIAAC-L, a number of data sets were produced and are available for scientific use. All accompanying documentation, including the "*Notes to the User*", technical reports, questionnaire documentation, and codebooks are accessible without restrictions and can be directly downloaded from the website of the Research Data Centre PIAAC at GESIS or the GESIS Data Archive.<sup>15</sup> Access to the data can be obtained by registering with the Research Data Centre PIAAC and signing a Data Use Agreement. Merging with PIAAC 2012 data (Data File ZA5845, current version 2.2.0, doi: [10.4232/1.12660](https://doi.org/10.4232/1.12660)) is permitted and can be facilitated by using the variable *seqid*. Across PIAAC-L data sets the ID variable *pnrfestid* can be used for merging.

The most actual version of the PIAAC-L scientific use database includes data sets of all three PIAAC-L waves (Data File ZA5989, version accessible on 14.12.2017: 3.0.0, doi: [10.4232/1.12925](https://doi.org/10.4232/1.12925)):

---

<sup>15</sup> <https://dbk.gesis.org/DBKSearch/SDESC2.asp?no=5989&tab=3&tdb=E&tdab=2&tdac=1> or <https://www.gesis.org/en/piaac/rdc/>

- ZA5989\_Persons\_14
  - Units: all PIAAC-L 2014 respondents (anchor persons and household members 18+ with participation in data collection 2014)
  - Content: data from person questionnaire, including derived variables
- ZA5989\_Household\_14
  - Units: all PIAAC-L 2014 households
  - Content: data from household questionnaire, including derived variables
- ZA5989\_Weights\_14
  - Units: anchor persons 2014
  - Content: weighting factors
- ZA5989\_Persons\_15
  - Units: all PIAAC-L 2015 respondents (anchor persons and partners living in the same household with participation in data collection 2015)
  - Content: questionnaire data, derived variables, cognitive assessment data, proficiency measures (plausible values for PIAAC literacy and numeracy; weighted likelihood estimates (WLE) for PIAAC literacy and numeracy assessed in 2015 and for NEPS reading and mathematics)
- ZA5989\_Weights\_15
  - Units: anchor persons 2015
  - Content: weighting factors
- ZA5989\_Persons\_16
  - Units: all PIAAC-L 2016 respondents (anchor persons and household members 18+ with participation in data collection 2016)
  - Content: data from person questionnaire, including derived variables
- ZA5989\_Cognit\_16
  - Units: all PIAAC-L 2016 respondents (anchor persons and household members 18+ with participation in data collection 2016)
  - Content: data from three short tests of cognitive ability
- ZA5989\_Household\_16
  - Units: all PIAAC-L 2016 households
  - Content: data from household questionnaire, including derived variables
- ZA5989\_Weights\_16
  - Units: anchor persons 2016
  - Content: weighting factors
- ZA5989\_Calendar
  - Units: all PIAAC-L 2014 and 2016 respondents (anchor persons and household members 18+ with participation in data collection 2014 and/or 2016)
  - Content: data from biographical calendar, spell data
- ZA5989\_Registry
  - Units: all persons ever registered in PIAAC-L
  - Content: basic information on participation in the different waves of data collection; this data set is incremental and was updated each wave
- ZA5989\_NumberSeries\_16
  - Units: pre-selected anchor persons in 2016
  - Content: data from Number Series Study

With PIAAC-L, a large body of additional information for the anchor persons and context information about the anchor persons' household and household members was collected. The data allows several research questions to be pursued. In addition to cross-sectional analyses where the focus is on data from one wave only, a variety of questions that need longitudinal data can be addressed using data from PIAAC 2012 and/or any of the PIAAC-L waves. The assessment of key competencies in wave 2 (with instruments from PIAAC and the NEPS) and cognitive skills in wave 3 (using short scales from the SOEP) as well as the data from the Number Series Study allows to investigate mechanisms of skill maintenance and development or assess correlations between the cognitive measures obtained from different tests. In the future, some anchor persons who have given their explicit consent will continue to be interviewed within the scope of the SOEP. The project PIAAC-L expired at the end of 2017.

## References

---

- Bohlender, A., & Glemser, A. (2016). *SOEP-IS 2014 – Methodenbericht zum Befragungsjahr 2014 des SOEP-Innovationssamples*. SOEP Survey Papers 339: Series B. Berlin: DIW/SOEP. Retrieved from [http://www.diw.de/de/diw\\_01.c.389326.de/publikationen\\_veranstaltungen/publikationen/soep\\_survey\\_papers/soep\\_survey\\_papers.html](http://www.diw.de/de/diw_01.c.389326.de/publikationen_veranstaltungen/publikationen/soep_survey_papers/soep_survey_papers.html)
- Burkhardt, L., Silbermann, T., & Bartsch, S. (2018). *Weighting in PIAAC-L 2016*. Manuscript in preparation.
- Carstensen, C. H., Gaasch, J. C., & Rothaug, E. (2018). *Scaling PIAAC-L cognitive data: technical report*. Manuscript in preparation.
- Engelhardt, L., & Goldhammer, F. (2018). *Number Series Study (DIPF): Technical Report*. GESIS Papers 2018|01. Cologne: GESIS.
- GESIS (2012). *PIAAC Deutschland 2012 – Hintergrundfragebogen*. Retrieved from <https://www.gesis.org/piaac/fdz/daten/nationale-scientific-use-files/>
- Heineck, G., & Anger, S. (2008). *The returns to cognitive abilities and personality traits in Germany*. SOEPPapers on Multidisciplinary Panel Data Research 124. Berlin: DIW/SOEP. Retrieved from [http://www.diw.de/de/diw\\_01.c.100410.de/publikationen\\_veranstaltungen/publikationen/soeppapers/soeppapers.html](http://www.diw.de/de/diw_01.c.100410.de/publikationen_veranstaltungen/publikationen/soeppapers/soeppapers.html)
- OECD. (2013a). *OECD skills outlook 2013: First results from the Survey of Adult Skills*. Paris: OECD.
- OECD. (2013b). *Technical report of the Survey of Adult Skills (PIAAC)*. Paris: OECD.
- Otto, W., Disch, K., Beitz, C., Gebhardt, S., Neuert, C., & Menold, N. (2015). *PIAAC-Longitudinal (PIAAC-L) 2016 – Hintergrundfragebogen. Kognitiver Pretest*. GESIS Projektbericht 2015|11. Mannheim: GESIS – Pretestlabor.
- Rammstedt, B. (2013). *Grundlegende Kompetenzen Erwachsener im internationalen Vergleich – Ergebnisse von PIAAC 2012*. Münster: Waxmann.
- Rammstedt, B., Martin, S., Zabal, A., Carstensen, C., & Schupp, J. (2017). The PIAAC Longitudinal Study in Germany – rationale and design. *Large-scale Assessments in Education*, 5(4). doi: [10.1186/s40536-017-0040-z](https://doi.org/10.1186/s40536-017-0040-z)
- Richter, D., Rohrer, J., Metzinger, M., Nestler, W., Weinhardt, M., & Schupp, J. (2017). *SOEP scales manual (updated for SOEP-Core v32.1)*. SOEP Survey Papers 423: Series C. Berlin: DIW/SOEP. Retrieved from [http://www.diw.de/de/diw\\_01.c.389326.de/publikationen\\_veranstaltungen/publikationen/soep\\_survey\\_papers/soep\\_survey\\_papers.html](http://www.diw.de/de/diw_01.c.389326.de/publikationen_veranstaltungen/publikationen/soep_survey_papers/soep_survey_papers.html)
- Schupp, J., Herrmann, S., Jaensch, P., & Lang, F. R. (2008). *Erfassung kognitiver Leistungspotentiale Erwachsener im Sozio-ökonomischen Panel (SOEP)*. Data Documentation No. 32. Berlin: DIW. Retrieved from <http://www.econstor.eu/handle/10419/129243>
- Steinacker, G., Wolfert, S., & Thümmel, K. (2017). *Durchführung der 3. Erhebungswelle von PIAAC-L (Kooperative längsschnittliche Weiterverfolgung der PIAAC-Studie in Deutschland)*. GESIS Papers 2017|27 Cologne: GESIS.

- TNS Infratest Sozialforschung (2011). *SOEP 2006 – Methodenbericht zum Befragungsjahr 2006 (Welle 23) des Sozio-oekonomischen Panels*. SOEP Survey Papers 61: Series B. Berlin: DIW/SOEP. Retrieved from [http://www.diw.de/de/diw\\_01.c.389326.de/publikationen\\_veranstaltungen/publikationen/soep\\_survey\\_papers/soep\\_survey\\_papers.html](http://www.diw.de/de/diw_01.c.389326.de/publikationen_veranstaltungen/publikationen/soep_survey_papers/soep_survey_papers.html)
- TNS Infratest Sozialforschung. (2014a). *SOEP 2014 – Erhebungsinstrumente 2014 (Welle 31) des Sozio-oekonomischen Panels: Lebenslauffragebogen, Altstichproben*. SOEP Survey Papers 237: Series A. Berlin: DIW/SOEP. Retrieved from [http://www.diw.de/de/diw\\_01.c.389326.de/publikationen\\_veranstaltungen/publikationen/soep\\_survey\\_papers/soep\\_survey\\_papers.html](http://www.diw.de/de/diw_01.c.389326.de/publikationen_veranstaltungen/publikationen/soep_survey_papers/soep_survey_papers.html)
- TNS Infratest Sozialforschung. (2014b). *SOEP 2014 – Erhebungsinstrumente 2014 (Welle 31) des Sozio-oekonomischen Panels: Personenfragebogen, Altstichproben*. SOEP Survey Papers 235: Series A. Berlin: DIW/SOEP. Retrieved from [http://www.diw.de/de/diw\\_01.c.389326.de/publikationen\\_veranstaltungen/publikationen/soep\\_survey\\_papers/soep\\_survey\\_papers.html](http://www.diw.de/de/diw_01.c.389326.de/publikationen_veranstaltungen/publikationen/soep_survey_papers/soep_survey_papers.html)
- TNS Infratest Sozialforschung. (2015a). *SOEP 2015 – Erhebungsinstrumente 2015 (Welle 32) des Sozio-oekonomischen Panels: Haushaltsfragebogen, Altstichproben*. SOEP Survey Papers 275: Series A. Berlin: DIW/SOEP. Retrieved from [http://www.diw.de/de/diw\\_01.c.389326.de/publikationen\\_veranstaltungen/publikationen/soep\\_survey\\_papers/soep\\_survey\\_papers.html](http://www.diw.de/de/diw_01.c.389326.de/publikationen_veranstaltungen/publikationen/soep_survey_papers/soep_survey_papers.html)
- TNS Infratest Sozialforschung. (2015b). *SOEP 2015 – Erhebungsinstrumente 2015 (Welle 32) des Sozio-oekonomischen Panels: Lebenslauffragebogen, Altstichproben*. SOEP Survey Papers 287: Series A. Berlin: DIW/SOEP. Retrieved from [http://www.diw.de/de/diw\\_01.c.389326.de/publikationen\\_veranstaltungen/publikationen/soep\\_survey\\_papers/soep\\_survey\\_papers.html](http://www.diw.de/de/diw_01.c.389326.de/publikationen_veranstaltungen/publikationen/soep_survey_papers/soep_survey_papers.html)
- TNS Infratest Sozialforschung. (2015c). *SOEP 2015 – Erhebungsinstrumente 2015 (Welle 32) des Sozio-oekonomischen Panels: Personenfragebogen, Altstichproben*. SOEP Survey Papers 274: Series A. Berlin: DIW/SOEP Retrieved from [http://www.diw.de/de/diw\\_01.c.389326.de/publikationen\\_veranstaltungen/publikationen/soep\\_survey\\_papers/soep\\_survey\\_papers.html](http://www.diw.de/de/diw_01.c.389326.de/publikationen_veranstaltungen/publikationen/soep_survey_papers/soep_survey_papers.html)
- Zabal, A., Martin, S., Massing, N., Ackermann, D., Helmschrott, S., Barkow, I., & Rammstedt, B. (2014). *PIAAC Germany 2012: Technical report*. Münster: Waxmann.
- Zabal, A., Martin, S., & Rammstedt, B. (2016). *PIAAC-L Data Collection 2014: Technical Report: Follow-Up to PIAAC Germany 2012*. GESIS Papers 2016|17. Cologne: GESIS.
- Zabal, A., Martin, S., & Rammstedt, B. (2017). *PIAAC-L Data Collection 2015: Technical Report*. GESIS Papers 2017|29. Cologne: GESIS.

## Annex: Questions From Original SOEP Instruments

The construction of the PIAAC-L 2016 *household questionnaire* was based on the SOEP household questionnaire from 2015. Due to cuts to the contents of the SOEP instrument for PIAAC-L, only a number of questions were kept (numbers refer to the question numbers in the corresponding SOEP questionnaire documentation from TNS Infratest Sozialforschung, 2015a): 2, 5, 6, 10, 11, 12, 23, 24, 25, 26, 34, 36, 37, 38, 39, 41, 42, 48, 49, 50, 51, 52, 70, 71, 72, 73, 77, 80, 82, and 83. For some of these questions, slight textual adaptations were implemented.

The PIAAC-L 2016 *person questionnaire* is an instrument, containing in parts adapted questions from different sources, such as the SOEP (2008, 2014, and 2015), PIAAC 2011/12, the Adult Education Survey (AES 2010), and new questions. The following Table A.1 shows the source for each question or block of questions and whether there was an adaptation to the question(s) or whether the items were newly developed. The numbers here refer to the question numbers given in the documentation of the person questionnaire for PIAAC-L 2016 (ZA5989\_fb\_Persons\_16.pdf):

Table A.1. Person Questionnaire Sources

Question No(s).	Source Questionnaire	Adap.
127, 128	SOEP Biography 2015 <sup>a</sup>	Y
2, 4, 5, 47, 48, 50 - 52F, 65 - 67C, 72 - 75, 77, 82 - 89, 92 - 95, 99 - 102, 104 - 108, 110 - 111A, 115 -122, 124A - 126, 140	SOEP Individual 2015 <sup>b</sup>	N
1, 6 - 6_3_ka, 49, 64, 64A, 70, 71, 76, 79 - 81, 90, 97, 98, 103, 109A - 109_13_b, 113A - 114B, F123_2, 138 - 139A_14	SOEP Individual 2015 <sup>b</sup>	Y
54 - 62, F123_1, 129 - 137A4_m	SOEP Biography 2014 <sup>c</sup>	N
53, 63	SOEP Biography 2014 <sup>c</sup>	Y
96	SOEP Individual 2014 <sup>d</sup>	N
7 - 9, 14, 18, 19	SOEP Individual 2014 <sup>d</sup>	Y
3	SOEP Big Five, PIAAC-L 2014 <sup>e</sup>	N
17, 23	SOEP Individual 2008 <sup>f</sup>	Y
16, 20, 24, 33, 35	PIAAC DE 2011/12 <sup>g</sup>	Y
11, 12, 13A - 13C, 15, 21, 22, 25, 29, 31, 32A - 32C, 34, 36, 41 - 45	AES 2010 <sup>h</sup>	Y
10, 26 - 28, 30, 37 - 40, 46, 78, 91a - 91i, 123A1 - 123A3, 141 - 143	New	N/A

Notes: Adap. = Adapted; Y = Yes; N = No; N/A = Not applicable.

<sup>a</sup> TNS Infratest Sozialforschung (2015b). <sup>b</sup> TNS Infratest Sozialforschung (2015c). <sup>c</sup> TNS Infratest Sozialforschung (2014a).

<sup>d</sup> TNS Infratest Sozialforschung (2014b). <sup>e</sup> ZA5989\_fb\_Persons\_14.pdf.

<sup>f</sup> [http://www.diw.de/documents/dokumentenarchiv/17/diw\\_01.c.85359.de/personen\\_2008.pdf](http://www.diw.de/documents/dokumentenarchiv/17/diw_01.c.85359.de/personen_2008.pdf) <sup>g</sup> GESIS (2012).

<sup>h</sup> <https://dbk.gesis.org/dbksearch/SDesc2.asp?DB=D&no=5074&tab=3&tdab=1&tdac=4>