Large-scale Assessments
in Education

RESEARCH   **Open Access**

On the comparability of adults with low literacy across LEO, PIAAC, and NEPS. Methodological considerations and empirical evidence

Tabea Durda[1]* , Britta Gauly[2], Klaus Buddeberg[3], Clemens M. Lechner[2] and Cordula Artelt[1]

*Correspondence: tabea.durda@lifbi.de
[1] LIfBi–Leibniz Institute for Educational Trajectories, Bamberg, Germany
Full list of author information is available at the end of the article

Abstract

Background: In Germany, three large-scale surveys–the Level One Study (LEO), the Programme for the International Assessment of Adult Competencies (PIAAC), and the National Educational Panel Study (NEPS)–provide complementary data on adults' literacy skills that can be harnessed to study adults with low literacy. To ensure that research on low-literate adults using these surveys arrives at valid and robust conclusions, it is imperative to ascertain the comparability of the three surveys' low-literacy samples. Towards that end, in the present study, we comprehensively assess the comparability of adults with low literacy across these surveys with regard to their sociodemographic and socioeconomic characteristics.

Methods: We used data from LEO, PIAAC, and NEPS. We identified features of the sample representation and measurement of (low) literacy as potential causes for variations in the low-literacy samples across the surveys. We then compared the low-literacy samples with regard to their sociodemographic and socioeconomic characteristics and performed logistic regressions to compare the relative importance of these characteristics as correlates of low literacy.

Results: The key insight our study provides is that–despite different sample representations and measurement approaches–the low-literacy samples in the three surveys are largely comparable in terms of their socioeconomic and sociodemographic characteristics. Although there were small differences between the surveys with regard to the distribution of gender, educational attainment, and the proportion of non-native speakers within the group of low-literate adults, results revealed that both the prevalence of low literacy and its correlates were largely robust across LEO, PIAAC, and NEPS. Across all three surveys, lower educational attainment emerged as the most significant correlate of low literacy, followed by a non-German language background, unemployment and low occupational status.

Conclusions: Our study provides evidence that all three surveys can be used for investigating adults with low literacy. The small differences between the low-literacy samples across the three surveys appear to be associated with sample representation and certain assessment features that should be kept in mind when using the surveys for research and policy purposes. Nevertheless, our study showed that we do not

compare apples with oranges when dealing with low-literate adults across different large-scale surveys.

Keywords: Low literacy, Adulthood, Survey research, Monitoring, LEO, PIAAC, NEPS

Introduction

The population of adults with low literacy in Germany comprises–depending on the survey used – 12.1 to 17.5 percent of the working-age population (Durda et al. 2020; Grotlüschen et al. 2019a; Grotlüschen and Riekmann 2011; OECD 2013b). Low-literate adults constitute a rather heterogeneous group of adults, for example, including older adults and adults with a non-German background who did not have the chance to acquire the necessary literacy skills, young adults with no educational qualification, employed adults who are trapped in low-skilled jobs, and the long-term unemployed. Often, their low literacy skills limit these adults' opportunities in the workplace as well as their access to health-related resources, social and political participation. This poses a challenge to policymakers, practitioners, and the wider society (Green 2013; Grotlüschen et al. 2016; Windisch 2015). Large-scale surveys play a key role in this context as they help relevant stakeholders to monitor trends in the prevalence of low literacy, to inform about the needs and resources of the target group, to influence educational policy, or to point out directions in adult education (e.g., Authoring Group Educational Reporting 2016).

In Germany, three large-scale surveys–all conducted at the outset of the past decade–provide complementary information on adults with low literacy with regard to those aspects: the Level One Study (LEO; Grotlüschen and Riekmann 2011), the Programme for the International Assessment of Adult Competencies (PIAAC; OECD 2013b), and the National Educational Panel Study (NEPS; Blossfeld and Roßbach 2019). The availability of these comprehensive surveys is a boon to policymakers, practitioners, and researchers since it gives them access to a richer source of information regarding adults' literacy skills as well as factors associated with the acquisition, retention, and maintenance of these skills. For example, NEPS offers insights into the malleability of low literacy owing to its longitudinal design (Wicht al. 2020). PIAAC allows international comparability in over 30 countries and provides detailed information on the relationship between literacy proficiency and employment including cognitive skills and tasks used at the workplace (OECD 2019a; Zabal et al. 2013), and its longitudinal follow-up in Germany (PIAAC-L) allows studying change in literacy over time (e.g., Gauly and Lechner 2019; Reder et al. 2020). Finally, LEO with its focus on adults with low literacy offers detailed insights into these adults' literacy practices and poses the question of social participation in depth (Grotlüschen et al. 2019b, 2020).

However, for research on low-literate adults utilizing these surveys to arrive at valid and robust conclusions, it is crucial to ensure that the subsamples of adults with low literacy in these three surveys are comparable. Ideally, comparability can be claimed if survey users could be reassured that adults identified as those with low literacy skills master the same proficiency levels and that they have the same sociodemographic and socioeconomic profiles. If the low-literacy samples are comparable, the surveys can be used for comparative analyses. There are, however, several potential threats to comparability that need to be considered before such a claim can

be made (Groves and Lyberg 2010). For example, if the test in one survey measures different aspects of literacy (e.g., writing, reading), or if some subgroups are over- or underrepresented in one low-literacy sample (e.g., non-native speakers, low-skilled adults), then the subgroup of adults with low literacy of one survey may not be comparable to that in another survey, or only to a limited extent. If this were the case, analysts using these surveys would need to know in which aspects the low-literacy samples differ, and which features of the sample representation and measurement of (low) literacy explain possible variations between the low-literacy samples.

Studies assessing the comparability of low-literacy adults in the three surveys are very sparse. Exceptions comprise, for example, Buddeberg et al. 2020, who conducted a linking study and showed that LEO and PIAAC do not measure identical but highly similar literacy constructs with a correlation of $r = 0.69$ and that the proficiency levels can be related to each other. Nevertheless, the findings that can be drawn from both surveys are not entirely consistent, for example, with LEO having relatively more men and non-native speakers in the low-literacy sample than PIAAC (Grotlüschen and Riekmann 2012; OECD 2013b), which means that currently, the survey user does not know where these differences come from. A comparison of the low-literacy sample of NEPS with LEO and PIAAC is completely missing so far.

In light of this research gap and its relevance to both educational research and educational policy, this paper aims to establish whether and to what extent the low-literacy samples across LEO, PIAAC, and NEPS are comparable with regard to key sociodemographic and socioeconomic characteristics.

We proceed as follows: First, we outline distinct features of the sample representation and measurement of (low) literacy as potential causes for variations in the low-literacy samples across LEO, PIAAC, and NEPS. We then outline relevant correlates of low literacy proficiency, before empirically examining differences in the distribution of relevant sociodemographic and socioeconomic characteristics in the sample of adults with low literacy and comparing the relative importance of those correlates for low literacy across all three surveys.

Thus, the present study aims to advance our knowledge about the comparability of the adults with low literacy with regard to their sociodemographic and socioeconomic characteristics, and about potential causes for variations in the low-literacy samples. It further adds to prior research on correlates of low literacy. Several studies have highlighted the role of socioeconomic factors, such as educational attainment or occupational status, and sociodemographic factors, such as gender, native language or age, and yet some findings have been inconsistent (e.g., Grotlüschen et al. 2012b; OECD 2013b). Moreover, prior research has not provided a conclusive understanding of the most important correlates of low literacy, because only some of those indicators in different surveys have been compared, studies have focused only on individual surveys, or the indicators have been operationalized in different ways. If we find convergence in the magnitude of those correlates for low literacy across LEO, PIAAC, and NEPS, then those findings can be taken as quite robust and compelling, given that the three surveys are conducted separately by different organizations, serve different purposes, and contain different assessments.

Potential causes for variations in the low-literacy samples

This section aims to explore potential causes for variations in the composition and coverage of the group of adults with low literacy. In line with the Total Survey Error approach (Groves and Lyberg 2010), we will distinguish two kinds of sources for variation: variation in sample representation and variation in measurement of (low) literacy.

Variation in sample representation

Variation in sample representation include all aspects that are related to the sampling. This includes in particular the definition of the target population, sampling procedure, response rates and non-response error. Table 1 gives an overview of the specifics of the sample representation.

Representation of target population

All surveys cover the same target population, that is, the working-age population in Germany, but they differ in the age ranges covered: 18 to 64 years (LEO), 16 to 65 years (PIAAC), and 24 to 69 years (NEPS). Further, what the surveys have in common is that all respondents who participated in the literacy assessment were interviewed via a computer-assisted personal interview (CAPI) at home in the language of their country of residence, and only respondents who spoke German sufficiently well were interviewed and tested (Bilger et al. 2012; Hammon et al. 2016; Mohadjer et al. 2013a).

Sampling procedure

For sampling, all three surveys employed complex sampling procedures to ensure the representativeness of the population, using a registry-based two-stage stratified and clustered random sampling design in NEPS and PIAAC (Hammon et al. 2016; Zabal et al. 2014) and a random route procedure in LEO (Bilger et al. 2012). The assessment of LEO took place as an add-on to the Adult Education Survey (AES). It used the regular AES-sample of 7,500 adults, and an additional sample of adults with low formal education was used to achieve a larger proportion of low-literate adults.

Response rates and accounting for non-response

Response rates and non-response bias are further features that can cause severe differences in the representation of the samples. If, for example, adults with low educational attainment are less likely to participate in and/or more likely to drop out from a panel, this could entail an underestimation of the proportion of adults with low literacy (e.g., Martin et al. 2020). Overall, all surveys under consideration here have acceptable response rates at the time of measuring literacy skills given that all surveys were voluntary. When interpreting the response rates, however, it must be noted that the response rate at the time of measuring literacy skills in NEPS refers to panelists and not to first respondents, thus it rather indicates respondent' willingness to continue participating in panel. If, on the other hand, we look at the initial response rate, the response rate for NEPS respondents is significantly lower, with response

Table 1 Comparing the Sample Representation and Measurement of (Low) Literacy of LEO, PIAAC, and NEPS

	LEO	PIAAC	NEPS[a]
Sample representation			
Target population	Residential population, living in private households in Germany, aged 18 to 64	Residential population, living in private households in Germany, aged 16 to 65	Residential population, living in private households in Germany, aged 24 to 69
Sampling design	Random route procedure, add-on to the AES	Registry-based two-stage stratified and clustered random sampling design	Registry-based two-stage stratified and clustered random sampling design
Sample size[b]	8,436	5,465	8,480
Response rate	52% (AES), 43% LEO (add-on)	55%	73% (wave 3)/77% (wave 5)[c]
Weighting	Design weights are adjusted to parameters of official statistics	Design weights are adjusted to parameters of official statistics and replicate weights (for more details on the exact weighting procedure in PIAAC, see OECD 2013a)	Design weights are adjusted to parameters of official statistics
Conceptual equivalence			
Focus	Reading and writing Lower end of the ability spectrum	Reading Entire ability spectrum	Reading Entire ability spectrum
Cognitive requirements		55% access and identify, 29% integrate and interpret, 16% evaluate and reflect	43% finding information in text, 27% drawing text-related conclusions, and 30% reflecting and assessing
Level	Letter, word, sentence and text	Text	Text
Medium	Audio and print	Digital and print	Print
Text contexts	Work related, personal	Work-related, personal, society and community, education, and training	Informational, instructional, advertising, commenting or argumenting, literacy
Text functions		Description, narration, exposition, argumentation, instruction, records	
Text format	Continuous, non-continuous	Continuous, non-continuous, and mixed texts; single and multiple texts	Continuous and single
Item format	Closed (writing and reading) Open-constructed (writing)	Closed Open-constructed (e.g. highlighting)	Closed

Table 1 (continued)

	LEO	PIAAC	NEPS[a]
Test design			
Delivery mode	Paper–Pencil-Assessment	Computer-Based-Assessment (default) or Paper–Pencil-Assessment	Paper–Pencil-Assessment
Test assembly	Booklet design with 72 items in total, 10 items for all respondents, additional 20 items out of three booklets for respondents who did not reach minimum number of correct answers	Adaptive testing; 2/3 of the respondents were administered 20 items in two stages from 58 items in total	One booklet for all respondents, 32 items in total per respondent
Administration	Background questionnaire was administered before assessment of competencies; No time restriction, on average 25 min	Background questionnaire was administered before assessment of competencies; No time restriction, on average 30 min	Background questionnaire was administered after assessment of competencies; 28-min time restriction
Scaling			
Scaling model	1-PL model	2-PL model	1-PL model
Missing responses	Not treated as incorrect	Some types of missing responses were treated as incorrect	Not treated as incorrect
Estimator	5 Plausible Values	10 Plausible Values	Warm's mean weighted likelihood estimation (WLE)
Standard setting			
Number of levels	5 proficiency levels with respondents at alpha-level 3 and below as low-literate	5 proficiency levels with respondents at level 1 and below as low-literate	2 proficiency levels with respondents at level 1 as low-literate and no further distinction for literates
Approach	A priori	Post hoc	A priori
Response probability	62%	67%	67%

[a] NEPS Starting Cohort Adults (SC6), wave 3 and 5.

[b] Note that these numbers refer to the number of participants who participated in the assessment.

[c] As the respective waves only includes panelists and no first respondents, it rather indicates respondent' willingness to continue participating in panel. The initial response rate for NEPS respondents is lower, with response rates around 30% (Hammon et al. 2016).

rates around 30% (Hammon et al. 2016). To address errors in the representation of the target populations, all surveys apply some kind of weighting procedure; all provide sampling weights to account for different selection probabilities (if applicable), to adjust for undercoverage and nonresponse, and are benchmarked to official statistics for central variables (Bilger et al. 2012; Hammon et al. 2016; Zabal et al. 2014). PIAAC also provides replicate weights to account for sampling variance (Zabal et al. 2014). With respect to the full samples, however, previous research shows that despite weighting adjustments, the distributions of the level of educational attainment, employment, and age groups between NEPS, PIAAC, and LEO slightly differ with reference to official statistics even though they are coded in a similar way (Bilger et al. 2012; Hammon et al. 2016; Widany et al. 2019; Zabal et al. 2014). In particular, it was shown that NEPS respondents are slightly older and the share of unemployed adults is significantly lower compared to the other surveys. Further, respondents of NEPS with a low and high educational attainment are relatively under- and overrepresented, respectively, compared to the AES (including parts of the LEO sample) and PIAAC.

Variation in measurement of (low) literacy

Variation in measurement includes all aspects related to the construct of interest (Groves and Lyberg 2010), namely, low literacy. This includes, in particular, the conceptual equivalence of the literacy assessments, the test design, the assessment procedure, the scaling procedure, and how the surveys account for measurement error, and the standard-setting procedure used to assign respondents to the low-literacy group. Table 1 gives an overview of the specifics in the measurement of (low) literacy.

Conceptual equivalence of the literacy assessments

The conceptual equivalence of the literacy assessments is probably the most decisive factor for variations in the low-literacy samples. The detailed comparison of the constructs reveals that all surveys assess literacy. They all understand literacy as an ability to understand and use written texts encouraging a person to achieve one's own goals, to develop one's own knowledge and potential, and to participate in social life (Gehrer et al. 2013; Grotlüschen et al. 2012; OECD 2013c; PIAAC Literacy Expert Group. 2009). However, the concrete assessment frameworks and coverage of the literacy constructs are different, which is most likely related to the respective objective. The primary difference is that LEO assesses not only reading but also writing competencies at the lower levels of literacy, whereas PIAAC and NEPS assess reading literacy on the entire ability spectrum. Therefore, the reading tasks in NEPS and PIAAC are designed in such a way that respondents must be able to read and retrieve information from sentences and text passages, i.e. a basic level of reading skills are assumed in the respondents (Gehrer et al. 2013; OECD 2013c; PIAAC Literacy Expert Group 2009), whereas in LEO, these basic skills are explicitly assessed at the level of letter, word, sentence, and text with stimulus material ranging from audio material to single words and short texts (Grotlüschen et al, 2012). It is therefore not surprising that a linking study has shown that the items used in LEO on average are much easier than the items used in PIAAC (Buddeberg et al. 2020); the same can be assumed for the NEPS-items. However, the results of the linking study also show that LEO and PIAAC measure highly similar literacy constructs with a

Durda *et al. Large-scale Assess Educ*   (2020) 8:13

Page 8 of 34

correlation of $r = 0.69$ and that the proficiency levels can be related to each other. The first preliminary results from a linking study between PIAAC and NEPS also show a high correlation of $r = 0.87$ between the literacy instruments (Carstensen et al. 2017).

As can be inferred from the focus of PIAAC and NEPS described above, despite notable differences in the assessments of reading literacy, there is larger conceptual overlap between the two assessments. PIAAC and NEPS both focus on comparable cognitive processes involved in the reading process. These cognitive processes represent different reading purposes including retrieval, integration, and evaluation of information and reflecting on form and content (Gehrer et al. 2013; OECD 2013c; PIAAC Literacy Expert Group 2009). In both assessments, the reading tasks are designed in such a way that these cognitive processes are easier or more difficult to accomplish, depending on the information contained in the text, for example, in terms of semantic match or distracting information (for further details on difficulty-generating factors see Durda et al. 2020). However, the constructs differ significantly in the types of the texts with which readers are confronted. PIAAC differentiates a variety of text types including different text formats, mediums, text functions, and text contexts (OECD 2013c; Yamamoto et al. 2012b). This distinction goes, therefore, beyond the text types covered by NEPS. To enable a coherent assessment of reading literacy over the lifespan, NEPS distinguishes between five different text functions rather than different texts contexts. In addition, compared to PIAAC, NEPS uses a narrower definition of reading literacy as the reading assessments contain only single, continuous, and printed texts (Gehrer et al. 2013). Given that the type of texts influence how readers process the text (e.g., the extent to which they need navigation skills through digital texts), it can be assumed that the two assessments place partly different cognitive and metacognitive demands (Barzillai et al. 2018; Solheim and Lundetræ 2018). Another difference that may affect the way a reader processes the texts and which other skills he or she requires concerns the item formats (e.g., Rauch and Hartig 2010; Solheim and Lundetræ 2018). Whereas PIAAC items contain both open-constructed response formats (e.g. highlighting) and closed response formats (e.g. multiple choice), the first cycle of NEPS used only closed item formats (Gehrer et al. 2013; OECD 2013c; Yamamoto et al. 2012b).

Test design

One major difference with regard to the test design occurs with respect to the delivery mode of the literacy assessments. While in LEO and NEPS,[1] the delivery mode was paper-based, PIAAC provided two modes: paper-based and computer-based. The default option for skill assessment in PIAAC was computer-based. However, if respondents had no or only very limited computer experience, refused to do the computer-based assessment or if they failed a very basic core test consisting of (easy) literacy and numeracy items, they completed a paper-based version. In Germany, 85 percent of respondents completed the computer-based version and 15 percent completed the paper-based version (Zabal et al. 2014). Further, the assessments differ with respect to the test assembly. The PIAAC assessment was based on a multi-matrix design with an adaptive multistage

---

[1] The delivery mode is paper-based at the first measurement point; computer-based assessment at the second measurement point.

testing in which two-third of the respondents were administered the literacy items, each consisting of a subset of 20 items in two stages from an item pool comprising a total of 58 literacy items. For those who merely worked on the other two skill domains of PIAAC (numeracy and problem-solving), literacy skills were imputed (Kirsch et al. 2013; Yamamoto et al. 2012a; Zabal et al. 2014). The LEO test was also conducted in a multi-matrix design such that each respondent was administered a subset of 10 items. In case the respondent did not reach a minimum number of correct answers, one out of three additional booklets with around 20 items was randomly administered. The entire item pool consisted of 72 items, all of them contained reading tasks, a part of the items contained also writing tasks (Grotlüschen et al. 2012; Hartig and Riekmann 2012). In contrast, all NEPS respondents received the same reading items in the same order comprising 32 reading items (Haberkorn et al. 2012; Koller et al. 2014).

Assessment procedure

Differences are also apparent with respect to the assessment procedure. In PIAAC and LEO, respondents were administered a background questionnaire before the assessment of competencies; in NEPS, a background questionnaire was administrated afterwards (Haberkorn et al. 2012; Grotlüschen et al. 2012; Koller et al. 2014; Zabal et al. 2014). Second, there was no time limit for test completing in PIAAC and LEO, but a 28-min time restriction in NEPS (Haberkorn et al. 2012; Koller et al. 2014). Nevertheless, the average time respondents spent on the literacy assessments are comparable across the surveys, ranging from 25 min in LEO (Grotlüschen et al. 2012a, b) to 30 min in PIAAC (Kirsch et al. 2013).

Scaling procedure and accounting for measurement error

Another source of variations between the surveys concerns the scaling procedure and the way the surveys take measurement error into account. First, differences occur in the choice of the item response model and the handling of missing data; both aspects can affect the estimation of a person's proficiency (e.g., Pohl et al. 2014; Robitzsch 2011; Rose et al. 2017). LEO and NEPS scaled the data assuming a 1-PL model, whereas PIAAC assumes a 2-PL model (Haberkorn et al. 2012; Hartig and Riekmann 2012; Koller et al. 2014; Pohl and Carstensen 2012; Yamamoto et al. 2012a). In NEPS and LEO, missing responses were not treated as incorrect responses; whereas, in PIAAC, some types of missing responses were treated as incorrect, for example, omitted items were treated as incorrect when the respondents spend more than five seconds on the item (Haberkorn et al. 2012; Hartig and Riekmann 2012; Koller et al. 2014; Pohl and Carstensen 2012; Yamamoto et al. 2012a). Further, differences occur in the way the surveys take measurement errors into account. This is important as measurement error bias the true test score of respondents (Braun and von Davier 2017), and thus can also affect the misclassification to the low-literacy group. To take the measurement error into account, PIAAC and LEO provide several literacy scores for respondents in the form of plausible values (Grotlüschen et al. 2012; Mohadjer et al. 2013b), for NEPS Warm's mean weighted likelihood estimation (WLE) are provided, showing good marginal reliabilities of 0.72, and 0.74 (Haberkorn et al. 2012; Koller et al. 2014).

Standard-setting procedure

Furthermore, assigning respondents to the low-literacy group also depends on the standard-setting procedure which divides the continuous latent scale into meaningful proficiency levels (Blömeke and Gustafsson 2017; Cizek 2012). In LEO, the decision on the number of proficiency levels and the description of each proficiency level was defined *a priori*. Proficiency level descriptors were used for item assignment to the respective proficiency level. The assignment of the items was revised for eight of 27 items after item calibration. In the final step, the mean of the item difficulties at each proficiency level was then used as the lower cut score of a proficiency level and the upper cut score representing the mean difficulty of the next higher proficiency level. There are five levels of proficiency (alpha-levels 1 to 5) with adults at alpha-level 3 and below classified as "low-literate" (Hartig and Riekmann 2012).

In PIAAC, proficiency is regarded as an ability continuum and displayed on a scale ranging from 0 to 500 points.[2] The scale can be divided into several proficiency levels. These levels are used for interpreting the proficiency results and are described in terms of the type of tasks adults with proficiency scores within a defined range are likely to complete successfully. There are five levels of literacy proficiency (below Level 1, Level 1 to 5) with adults at Level 1 and below classified as "low-literate" (for a detailed description of the Levels see OECD 2013b).

Due to its longitudinal design and related constraints, the provision of proficiency levels across the whole ability range was not intended for the NEPS assessments. However, a proficiency level approach was applied afterwards in order to be able to define low literacy and to investigate the causes for and changes among adults with low literacy. The standard-setting procedure thereby followed an a priori theory-driven approach. In order to differentiate between low-literate and literate adults, the Bookmark method was used. In this method, the reading items were presented in an "Ordered Item Booklet" (OIB) in which they were arranged according to item difficulty as determined by the applied Item Response Theory scaling model (Haberkorn et al. 2012; Koller et al. 2014), beginning with the easiest item. It was then the task of the panelists, through repeated comparison of the reading items with proficiency level descriptors, to set the cut score "bookmark" between those items that, in their view, define the boundary between a low reading proficiency level and a functional level of reading proficiency (Durda et al. 2020). Because the Bookmark procedure is not free of criticism regarding its validity (see for review Karantonis and Sireci 2006; Lin 2006), the cut score was cross-validated with a mixture Rasch model in a recent study. Results suggested a high agreement of almost 90% in the proficiency assignment of the respondents to the low-literacy group between the Bookmark procedure and the mixture Rasch model.

In this context, the choice of the response probability is an important part of the standard-setting procedure as it can influence the location of the cut score and thus the percentage of adults that fall into the low-literacy group. In LEO, the response probability was set at 62 percent and in PIAAC and NEPS at 67 percent (Durda et al. 2020; Hartig and Riekmann 2012; Yamamoto et al. 2012b).

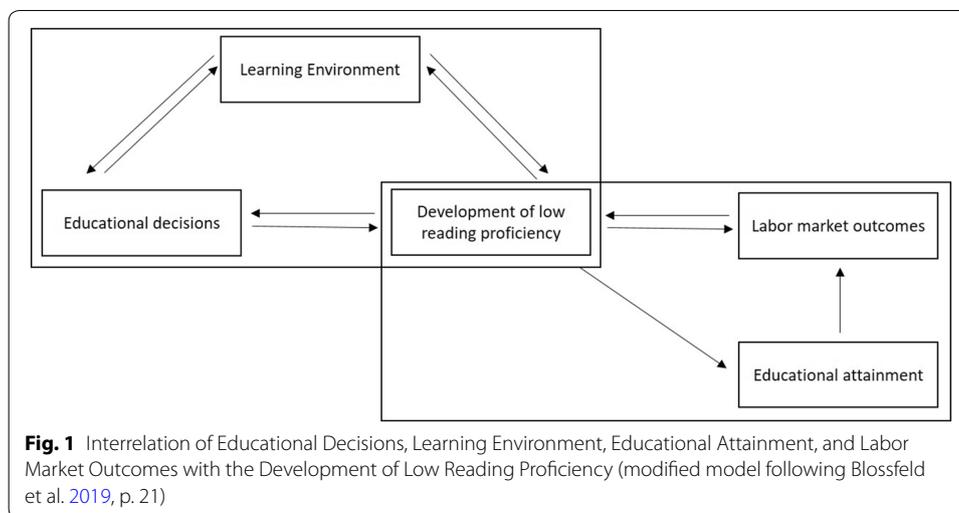[2] In theory, the scales range from $-\infty$ to $+\infty$.

Based on these procedures, low-literate adults are described as follows:

– LEO: Alpha-level 3 (…) is used when people can read or write single sentences, but not continuous text—even if it is brief. Due to their limited written language skills, people affected cannot properly deal with the everyday requirements of life in society. For example, even if they do simple jobs they cannot read written instructions (Grotlüschen and Riekmann 2011, p. 6).
– PIAAC: Adults at Level 1 and below can read and locate a single piece of information in relatively simply texts in which the information requested is identical or synonymous with the information given in the question or directive. Adults at this proficiency level usually understand sentences or sections but are generally unable to process, compare, and evaluate several pieces of information (OECD 2013b, p. 66f).
– NEPS: Adults at Level 1 can locate single pieces of information and can cycle through more than one piece of information between neighboring sentences (local inferences). The required information is in most parts literal with the information in the task. The tasks contain minimal distracting information (if there is distracting information, it occurs after the solution). Solving the task does not require detailed comprehension of the task or text but basic understanding. Often adults can use text-signaling devices to locate the required information (Durda et al. 2020).

Summary

The foregoing comparison made it clear that the surveys are similar in several characteristics but differ in some others. With regard to the comparability of the sample representation, the most striking differences occurred with respect to the age range and response rates of adults with low educational attainment and unemployed adults. While variation in age range can be countered by limiting analyses to the same age range, variation in the coverage of educational attainment and unemployed adults is more challenging. Due to a lower proportion of unemployed adults and low-qualified adults in the full sample of NEPS, it can be assumed that these respondents will also be comparatively less represented in the low-literacy sample compared to LEO and PIAAC.

With regard to the comparability from the perspective of measurement, the findings, so far, suggest that the assessments capture different facets of (low) literacy and therefore, are not fully congruent, even though they largely capture the same construct (i.e. literacy), and that the proficiency levels can be related to each other. However, because of the different operationalizations, scaling and standard-setting procedures, we would expect to see variations within the corresponding low-literacy samples as well. For example, previous studies have shown that non-continuous texts have advantages for men (Solheim and Lundetræ 2018). Consequently, it is to be expected that men—even in the low-literacy group—perform better in PIAAC than, for example, in NEPS. Similarly, the proportion of native and non-native speakers in LEO is expected to differ compared to NEPS and PIAAC, since non-native speakers perform probably better in a literacy test that measures reading only compared to a literacy test that measures reading and writing (Grotlüschen et al. 2012).

Fig. 1 Interrelation of Educational Decisions, Learning Environment, Educational Attainment, and Labor Market Outcomes with the Development of Low Reading Proficiency (modified model following Blossfeld et al. 2019, p. 21)

In sum, this section described a range of characteristics at the level of sample representation and measurement as potential causes for variations in the low-literacy samples across LEO, PIAAC, and NEPS. Although our study does not base upon a common scaling of all three surveys, it is now possible to give a clearer picture of the factors that can explain variations in the low-literacy samples.

Correlates of low literacy in adulthood

In cross-sectional studies, various factors are typically used to characterize low-literate adults such as sociodemographic characteristics (e.g. age, gender) and factors describing their socioeconomic situation (e.g., Grotlüschen et al. 2016). Following a life course perspective, low literacy is strongly interrelated with formal and non-formal learning environments, educational decisions, and educational and labor market outcomes (see Fig. 1). Although our cross-sectional comparison does not empirically capture the life course dynamics one can make well-grounded theoretical assumptions about plausible associations between sociodemographic and socioeconomic stations and (low) literacy during adulthood.

Socioeconomic correlates of low literacy in adulthood

Several cross-sectional studies indicate a close relationship between *educational attainment* and literacy. In this context, the first assumption we make is that educational attainment is an important correlate of low-literacy as it can partly be seen as the result of an unfavorable socialization in literacy, depending on non-formal and formal learning opportunities, and partly as a cause of educational decisions that lead to lower reading performance and thereby to lower educational attainment. A person's reading literacy development does not start with the beginning of formal instruction in school but is influenced by experiences gained in early childhood. It is well established that literacy-related practices before and during school are critical for the development of children's reading literacy skills, but literacy practices vary depending on families' social background. It has been shown that children raised in socially disadvantages families acquired pre-reading skills less often than necessary because their families engaged

them less often in literacy practices that were quantitatively and qualitatively adequate in terms of dominant literacy practices (e.g., Bus et al. 1995; Mol and Bus 2011). These disparities are not only present at primary school entry but persist throughout primary school. After primary school, most students in Germany are assigned one of up to three different tracks[3] of secondary school according to their abilities but, inadvertently, and also often according to the social background (e.g., Cortina et al. 2005; Ditton and Krüsken 2010): lower secondary school (Hauptschule), intermediate secondary school (Realschule) or upper academic school track (Gymnasium). Students in lower academic school track start on average with significantly lower reading skills in the beginning of secondary school than students at middle and upper academic tracks, and these differences often persist throughout secondary school or even extend (e.g., Pfost and Artelt 2014), resulting in a lower educational attainment.

Another strand of research has highlighted the relationship between literacy proficiency, educational attainment and labor market outcomes, for example, in terms of a *person's occupational status* or *employment status* (Arendt et al. 2008; Barone and Werfhorst 2011; Calero and Choi 2017; Grotlüschen et al. 2016; McIntosh 2001; OECD 2013b; Perry and Gauly 2019; Shomos 2010; Wicht et al. 2019). Undoubtedly, educational attainment influences opportunities for access to the labor market (Arrow 1973; Spence 1973), and this applies particularly to Germany with its strong emphasis on vocational training and formal educational credentials (Solga et al. 2014). However, because most occupations require the ability to engage with written materials, literacy proficiency also has a direct effect on labor market outcomes. For example, information such as work instructions are difficult to obtain without a functional level of literacy. Consequently, low-literate adults are often employed in occupations with a low socioeconomic status requiring little engagement with written language. These are, for example, semi-skilled and elementary occupations in the cleaning or construction sector (Grotlüschen et al. 2016; OECD 2013b). However, given that different occupations provide certain literacy environments, low literacy can also be the consequence of job requirements and characteristics. Practice engagement theory (Reder 2009) assigns literacy practices a key role in explaining literacy development during adulthood. For example, adults who find themselves more often in semi-skilled or elementary occupations are not only confronted less often with written material, but the material is also often less demanding. In this context, Smith (2000) showed that adults working in semi-skilled or elementary occupations read more often for functional purposes (e.g. reading to understand work instructions) that required less demanding reading techniques (e.g. note taking) compared to adults working in occupations that required more demanding reading techniques that help summarize, compare, and evaluate information. Further, longitudinal studies showed that lower work-related engagement with written material is associated with losses in reading literacy (Wicht et al. 2020; Bynner and Parsons 2006; Reder 2009). Both mechanisms explain why low-literate adults are more likely to have a lower occupational status.

Because educational attainment and labor market outcomes are interrelated, it is not surprising that *unemployment* rate is higher among adults with lower proficiency

---

[3] The school types in secondary school can differ depending on the federal state of Germany.

(Arendt et al. 2008; Calero and Choi 2017; Grotlüschen et al. 2016). Yet not only lower educational attainment but also literacy proficiency itself cause problems in (re-)entering the labor market (Arendt et al. 2008; Calero and Choi 2017). On the other hand, following practice engagement theory (Reder 2009), unemployment can also entail literacy losses by reducing engagement with written material due to lack of such requirements. In this context, the study "Use It or Lose It" (Bynner and Parsons 1998) showed that unemployment had negative long-term effects on reading proficiency especially for adults with poor initial reading literacy skills (but see Wicht al. 2020).

Further sociodemographic correlates of low literacy

In addition to educational attainment and labor market outcomes, cross-sectional studies have linked a range of further sociodemographic characteristics to low literacy proficiency (e.g., Grotlüschen et al. 2012a; OECD 2013b). These include in particular gender, native language, and age.

With regard to *gender*, consensus exist that females outperform males in literacy proficiency up to secondary school (Mullis et al. 2017; OECD 2019b), whereas findings for adulthood are inconclusive (Grotlüschen et al. 2012b; OECD 2013b; Thums et al. 2020). Whereas some studies indicate that males somewhat catch up and gender differences disappear during adulthood (OECD 2013b; Thums et al. 2020), other studies point to an overrepresentation of men among low-literates (Grotlüschen and Riekmann 2012). Various explanations have been offered for inconsistent findings, ranging from differences in reading motivation and behavior between male low-literate and literate readers (e.g., in terms of resistance to and avoidance of literacy practices; Frijters et al. 2019; Lenters 2006), the role educational inequalities and systems (e.g., in terms of an overrepresentation of male students in lower secondary education; Buchmann et al. 2008; Van Hek et al. 2019), or the nature of the literacy assessment that can affect male and female differently (e.g. non-continuous tests or continuous texts; digital reading or paper-based assessment; Barzillai et al. 2018; Solheim and Lundetræ 2018).

Further, large-scale surveys have also found that the proportion of non-native speakers is considerably higher at the lowest proficiency level and the effect persists after controlling for social background. The remaining effect can therefore, cautiously interpreted, be attributed to a poorer knowledge of the German language which causes lower literacy proficiency (Grotlüschen and Riekmann 2012, Grotlüschen et al. 2016; OECD 2013b).

Finally, low literacy tends to be more prevalent in older compared to younger adults (e.g., Desjardins and Warnke 2012; Flisi et al. 2019; Paccagnella 2016). These age differences in the prevalence of low literacy could reflect cohort differences in educational opportunities (e.g., the average years of schooling increased in younger cohorts). Alternatively, these age differences could reflect true age effects. The general slowing hypothesis (Choi and Feng 2015) attributes age-related declines in literacy during adulthood to a general loss in mental capacities that is due to biological ageing processes. In line with this idea, recent longitudinal findings suggested that age-related declines in reasoning (fluid intelligence) and perceptual speed are responsible for literacy losses among older adults (Wicht et al. 2020). Both mechanisms are not mutually exclusive and might explain jointly why older adults have lower literacy competencies.

Empirical comparison of low-literate adults across LEO, PIAAC, and NEPS

The previous sections described a range of characteristics that differ between the surveys. Further, we outlined relevant indicators for low reading proficiency: sociodemographic characteristics (in particular gender, native language, and age), and socioeconomic characteristics (particularly educational attainment, occupational status, and employment status). Next, we empirically compare the distribution of these characteristics among low-literate adults and their relative association with low literacy across LEO, PIAAC, and NEPS.

Method

Sample and measures

As stated before, we use representative samples of the German working-age population from a similar reference period (years 2010 to 2013) for all following analyses. The first cycle of LEO (Grotlüschen and Riekmann 2011) was carried out in 2010. The assessment took place as an add-on to the AES. In total, 8,436 interviews were carried out. The German sample of the first cycle of PIAAC (OECD 2013a, 2013b) comprised 5,465 interviews and was conducted in 2012. For NEPS, we used data from Starting Cohort 6 "Adults" (Blossfeld and Roßbach 2019) with two sub-samples with either a six-year retest interval ($N = 5,335$, initial assessment in wave 3, 2010/11) or a four-year interval ($N = 3,145$, initial assessment in wave 5, 2012/13), applying identical tests (Haberkorn et al. 2012; Koller et al. 2014). To ensure the comparability between the surveys, we restricted the age range between 24 to 64 years, which is the same for all.

Low Literacy. Low literacy is dummy coded, representing either that an adult belongs to the literacy group (0) or low-literacy group (1). Data on low literacy were obtained from the test scores of the respective survey. All data were scaled using Item Response Theory (IRT) with five or ten plausible values (PV) for LEO and PIAAC, respectively, and one Warm's mean weighted likelihood estimation (WLE) for each NEPS respondent. LEO respondents who scored at alpha-level 3 or below were assigned to the low-literacy group and those who scored at alpha-level 4 and above were assigned to the literacy group. Due to the plausible values approach, this results in five dummy coded variables for each respondent (Grotlüschen et al. 2012). In PIAAC, adults who scored at Level 1 or below were assigned to the low-literacy group, and those who scored at Level 2 and above were assigned to the literacy group, resulting in 10 dummy coded variables for each respondent (Mohadjer et al. 2013b). In NEPS, adults who scored at proficiency level 1 were assigned to the low-literacy group, and those who scored at proficiency level 2 were assigned to the literacy group, resulting in one dummy coded variable for each person (Durda et al. 2020; Haberkorn et al. 2012; Koller et al. 2014).

We selected the following sociodemographic and socioeconomic characteristics with a high comparability over all surveys:

Gender. Gender indicates whether the adult identifies as female (0) or male (1).

Age. For comparing the distribution, age was used as a continuous variable; for determining the relative importance, age was categorized into age bands of 10 years: (1) 24 to 33 years, (2) 34 to 43 years, (3) 44 to 53 years, and (4) 54 to 64 years.

Native Language. Native language indicates the first language spoken at home during childhood, representing the language background of the person with (0) German or (1) non-German.

Educational Attainment. Data on the level of educational attainment were obtained from information on a person's highest educational qualification and mapped onto the International Standard Classification of Education (ISCED-97, UNESCO Institute for Statistics 2006). We distinguish four levels of educational attainment: (1) high (ISCED 5/6), (2) medium (ISCED 3/4), (3) low (ISCED 2) and (4) no (ISCED 1 and below).

Occupational Status. Data on the occupational status were obtained from information on a person's occupational situation which was mapped onto the International Standard Classification of Occupations (ISCO-08), and then converted to ISEI-08 codes (International Socioeconomic Index of Occupational Status, Ganzeboom 2010). For determining the relative importance, the occupational status was categorized into four status groups: (1) very high (69.61 to 88.96), (2) high (50.26 to 69.60), (3) medium (30.92 to 50.25), and (2) low (11.56 to 30.91).

Employment Status. Employment status is indicated by three dummy variables, representing either adults in self-reported subjective status of (1) employment, (2) unemployment or (3) other employment status (e.g. maternal leave).

A detailed overview of all variables used in our analyses can be found in the Appendix in Table 4.

Analyses

We applied a two-fold strategy to compare the low-literate adults between the surveys. First, we compared the low-literate adults with regard to their sociodemographic and socioeconomic compositions. We conducted Chi-square tests for categorical variables (i.e., gender, native language, educational attainment, and employment status) and one-way ANOVAs for continuous ones (i.e., occupational status and age). Due to the large sample sizes, we focused on effect sizes (Phi ϕ, Cramer's V, Cohen's f) rather than statistical significance (see Appendix, Table 5 for interpretation). For estimating differences in the distribution only, we applied the survey-specific weighting factors allowing to adjust the sampling design for bias effect (Bilger et al. 2012; Hammon et al. 2016; Mohadjer et al. 2013b; see Table 4).

Second, we conducted two logistic regressions for each of the three surveys separately and calculated Average Marginal Effects (AMEs) to gauge associations between important sociodemographic and socioeconomic factors and low literacy proficiency (0 = literate, 1 = low-literate).
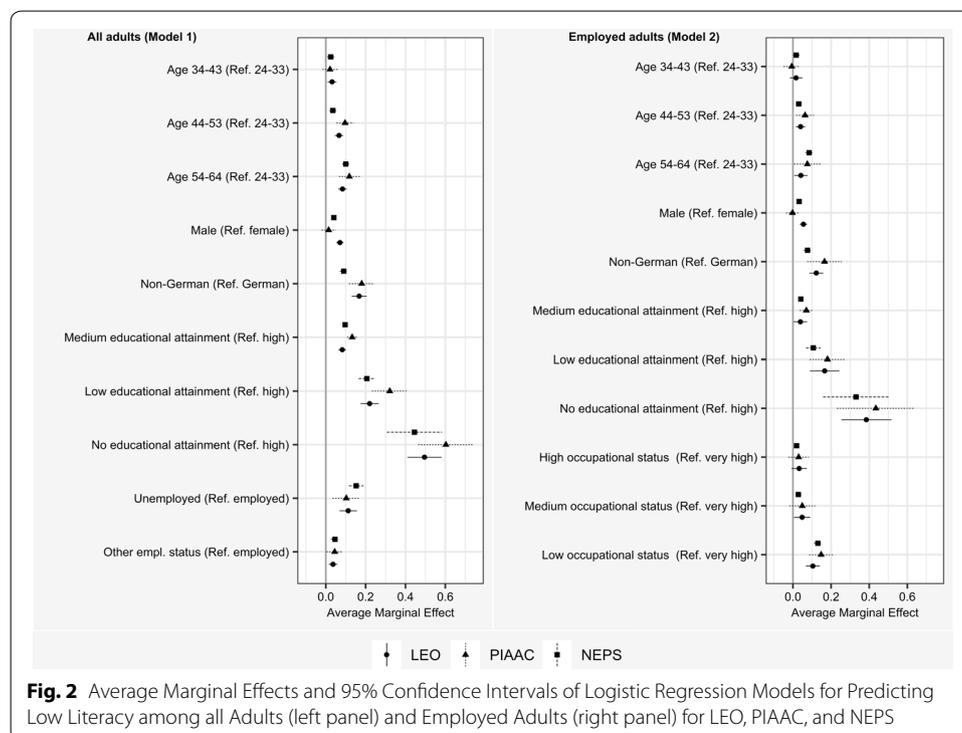
In Model 1, a logistic regression was performed for all adults including five correlates: age (grouped), gender, native language, educational attainment, and employment status. In Model 2, a logistic regression was run for employed adults only, including age (grouped), gender, native language, educational attainment, and occupational status (grouped). The respective reference groups are the following. The reference group in the age variable is youngest age group, adults between 24 to 33 years

of age. The reference group for men and adults with non-German language background are women and adults with a German language background, respectively. For the educational attainment, the reference group are adults with a high educational attainment (ISECD 5 and 6). For the employment status, the reference group are adults reporting to be in employment, and for the occupational status, the reference group are adults with the highest ISEI (69.61 to 88.96). We plotted the AMEs for each correlate in each of the surveys in Fig. 2 (see Tables 6, 7 in appendix for more details). Therefore, we could directly compare how the differences in the probabilities in the low-literacy group vary between different sociodemographic and socioeconomic factors and surveys.

The analyses were restricted to cases without missing values. With exception of occupational status for NEPS, the share of missing values per variable is < 1.5% of the full sample of respondents aged 24 to 64 years across the surveys (Cheema 2014; Graham et al. 2012; see Table 8 in the appendix for more details).

Results

Our empirical comparison of the surveys starts with an overview of sample characteristics of the full sample. Subsequently, we focus on the comparison of the distribution of the sociodemographic and socioeconomic factors among the subgroups of low-literate adults only. Finally, the estimated logistic regression models show if and how the association between the correlates and low literacy among all adults and employed adults vary across the three surveys.

Fig. 2 Average Marginal Effects and 95% Confidence Intervals of Logistic Regression Models for Predicting Low Literacy among all Adults (left panel) and Employed Adults (right panel) for LEO, PIAAC, and NEPS

Sociodemographic and socioeconomic characteristics for the full sample

Table 2 provides the sample statistics across LEO, PIAAC, and NEPS. A comparison of the full sample statistics reveals equal distributions for the share of low-literate adults in the sample, $X^2(2) = 34.72$, $p < 0.001$, $V = 0.05$, gender, $X^2(2) = 7.49$, $p = 0.024$, $V = 0.02$, for native language, $X^2(2) = 21.72$, $p < 0.001$, $V = 0.03$, and occupational status, F(2, 19,403) = 14.96, p < 0.001, $f = 0.04$. There are small differences in the age distribution, F(2, 19,403) = 118.6, p < 0.001, $f = 0.11$, suggesting that the respondents of NEPS are slightly older compared to those of PIAAC and LEO. Educational attainments also varies slightly across surveys, $X^2(6) = 290.53$, $p < 0.001$, $V = 0.09$, which is particularly evident in the share of adults with no and high educational attainment, with a lower representation of adults with no educational attainment and a larger representation of adults with a high educational attainment in the NEPS sample. Further, there are small differences with respect to employment status, $X^2(4) = 218.29$, $p < 0.001$, $V = 0.07$, with a slightly higher percentage of adults in employment for the NEPS sample.

Sociodemographic and socioeconomic characteristics for the low-literacy sample

As can be seen from Table 3, for all sociodemographic and socioeconomic characteristics except for occupational status, we found small differences in the distribution between the surveys among low-literate adults. The most pronounced differences emerged for native language and educational attainment.

We found small differences in the age distribution of low-literate adults in NEPS compared to LEO and compared to PIAAC, with on average older respondents in NEPS. For gender, there was a small difference between PIAAC compared to LEO and compared to NEPS with a higher proportion of male low-literate respondents in LEO (with a difference of 12.25%) and NEPS (with a difference of 11.29%) compared to PIAAC (see Table 2). Furthermore, the proportion of non-native speakers was higher in LEO compared to PIAAC (with a difference of 15.19%) and compared to NEPS (with a difference of 16.86%), however, the effect sizes were small.

The levels of educational attainment varied between the surveys, which was most pronounced in the proportion of low-literate adults with no educational attainment. In more detail, Chi-square tests indicated small differences in the proportion of low-literate adults with no educational attainment in NEPS compared to LEO, ranging from ϕ 0.19 to ϕ 0.31, and compared to PIAAC, ranging from ϕ 0.14 to ϕ 0.29, with highest differences for no educational attainment vs. high educational attainment. While 14.41% of low-literate adults in LEO and 12.52% of low-literate adults in PIAAC have no educational attainment, only 5.01% of NEPS respondents have no educational attainment. On the other hand, 14.63% of low-literate adults in NEPS have a high educational attainment compared with 10.75% in PIAAC and 11.23% in LEO. Effect sizes, however, were negligible for differences in the proportion of low-literate adults with low education vs. medium education and medium vs. high education between LEO, PIAAC, and NEPS.

Moreover, small differences in the distribution of employment status among low-literate adults emerged between LEO and PIAAC, LEO and NEPS, and PIAAC and NEPS. These differences occurred primarily for adults in unemployment and in another employment status (e.g. maternal leave) but not for the proportion of adults in

Table 2 Basic statistics of the analytical sample

| | Full sample | | | | Low-literate sample | | | | Literate sample | | | |
| | 95% | | | | 95% | | | | 95% | | | |
	%/M (SD)	SE	LL	UL	%/M (SD)	SE	LL	UL	%/M (SD)	SE	LL	UL
Literacy in %												
LEO	100.00	–	–	–	14.69	0.57	13.57	15.81	85.31	0.57	84.19	86.43
PIAAC	100.00	–	–	–	17.65	0.84	15.96	19.35	82.35	0.84	80.65	84.04
NEPS	100.00	–	–	–	13.69	0.64	12.49	15.00	86.30	0.64	85.00	87.51
Male in %												
LEO	50.19	0.71	48.78	51.59	60.03	1.99	56.11	63.96	48.49	0.79	46.93	50.04
PIAAC	50.51	0.20	50.11	50.90	47.78	2.53	42.52	53.04	51.09	0.58	49.90	52.28
NEPS	52.30	0.85	50.63	53.96	59.07	2.51	54.07	63.89	51.23	0.90	49.46	52.99
Age in M (SD)												
LEO	44.07 (11.04)	0.15	43.77	44.37	44.73 (11.02)	0.42	43.89	45.56	43.96 (11.04)	0.17	43.63	44.28
PIAAC	44.44 (11.10)	0.08	44.28	44.59	46.86 (10.89)	0.53	45.78	47.94	43.91 (11.07)	0.13	43.64	44.19
NEPS	46.66 (10.68)	0.18	46.30	47.02	48.98 (9.95)	0.61	47.78	50.17	46.29 (10.72)	0.18	45.92	46.66
Non-German in %												
LEO	14.85	0.58	13.72	15.98	42.67	3.02	36.08	49.25	10.06	0.64	8.76	11.36
PIAAC	12.43	0.56	11.31	13.55	27.48	2.06	23.34	31.61	9.20	0.58	8.04	10.36
NEPS	12.55	0.70	11.24	13.99	25.81	2.51	21.19	31.04	10.44	0.69	9.16	11.88
No edu. attainment (ISCED 1 and below) in %												
LEO	3.08	0.23	2.62	3.53	14.41	1.38	11.69	17.14	1.13	0.15	0.82	1.43
PIAAC	3.03	0.32	2.39	3.67	12.52	1.60	9.33	15.72	0.99	0.24	0.51	1.48
NEPS	1.08	0.23	0.72	1.63	5.01	1.29	3.01	8.23	0.46	0.16	0.23	0.89
Low edu. attainment (ISCED 2) in %												
LEO	9.18	0.39	8.42	9.94	21.56	1.82	17.90	25.22	7.05	0.39	6.27	7.82
PIAAC	6.94	0.46	6.02	7.87	17.73	2.12	13.44	22.03	4.63	0.47	3.69	5.57
NEPS	9.59	0.63	8.43	10.91	22.86	2.51	18.30	28.16	7.49	0.59	6.41	8.73

Table 2 (continued)

	Full sample				Low-literate sample				Literate sample			
	%/M (SD)	95%			%/M (SD)	95%			%/M (SD)	95%		
		SE	LL	UL		SE	LL	UL		SE	LL	UL
Medium edu. attainment (ISCED 3/4) in %												
LEO	59.91	0.71	58.52	61.29	52.80	2.07	48.71	56.89	61.13	0.77	59.62	62.65
PIAAC	55.66	0.80	54.07	57.25	59.00	2.49	54.00	63.99	54.95	0.79	53.37	56.52
NEPS	50.76	0.85	49.09	52.43	57.50	2.62	52.28	62.55	49.69	0.90	47.93	51.45
High edu. attainment (ISCED 5/6) in %												
LEO	27.84	0.66	26.54	29.14	11.23	1.70	7.73	14.73	30.69	0.75	29.22	32.17
PIAAC	34.37	0.66	33.05	35.68	10.75	1.31	8.13	13.37	39.43	0.75	37.93	40.93
NEPS	38.56	0.81	36.99	40.16	14.63	1.74	11.53	18.37	42.36	0.87	40.66	44.08
Employed in %												
LEO	71.68	0.58	70.54	72.83	59.58	2.08	55.41	63.75	73.77	0.64	72.52	75.01
PIAAC	74.07	0.73	72.61	75.53	60.44	2.26	55.90	64.98	76.99	0.84	75.31	78.67
NEPS	80.55	0.70	79.14	81.88	63.99	2.46	59.03	68.67	83.18	0.70	81.76	84.50
Unemployed in %												
LEO	7.66	0.31	7.06	8.27	17.17	1.54	14.02	20.32	6.03	0.32	5.39	6.67
PIAAC	5.48	0.40	4.68	6.29	10.29	1.48	7.33	13.24	4.45	0.42	3.61	5.30
NEPS	6.59	0.48	5.71	7.59	19.36	2.12	15.53	23.87	4.56	0.42	3.80	5.47
Other employ. status in %												
LEO	20.65	0.52	19.64	21.67	23.25	1.66	19.92	26.58	20.21	0.57	19.08	21.33
PIAAC	20.45	0.75	18.96	21.93	29.27	2.16	24.94	33.60	18.55	0.83	16.90	20.21
NEPS	12.86	0.57	11.78	14.02	16.65	1.79	13.42	20.47	12.26	0.60	11.14	13.48
Occupational status (ISEI-08) in M (SD)												
LEO	46.17 (20.73)	0.38	45.44	46.91	32.51 (16.71)	1.07	30.37	34.65	48.07 (20.52)	0.39	47.30	48.84
PIAAC	44.88 (20.91)	0.30	44.28	45.48	30.75 (14.35)	0.91	28.88	32.62	47.24 (20.90)	0.35	46.54	47.94
NEPS	47.05 (21.04)	0.39	46.29	47.82	30.90 (14.21)	0.99	29.05	32.92	49.02 (20.71)	0.40	48.23	49.80

Table shows the weighted distributions of the sociodemographic and socioeconomic characteristics in our three samples of respondents aged from 24 to 64 years across the surveys. The full analytic samples comprise of 7,438 (LEO), 4,335 (PIAAC), and 7,633 (NEPS) respondents. LL and UL indicate the lower and upper limits of a confidence interval, respectively

Durda *et al. Large-scale Assess Educ* (2020) 8:13

Page 21 of 34

Table 3 Summary of Independent Sample t-Tests for Age and Occupational Status and Chi-square Tests for Gender, Native Language, Educational Attainment, and Employment Status for LEO vs. PIAAC, LEO vs. NEPS, and PIAAC vs. NEPS among Low–Literate Adults

	LEO vs. PIAAC				LEO vs. NEPS				PIAAC vs. NEPS			
	t/X^2	df	p	ES	t/X^2	df	p	ES	t/X^2	df	p	ES
Gender	26.83	1	<0.001	*0.12*	0.17	1	0.682	0.01	22.26	1	<0.001	*0.11*
Age	4.12	1856	<0.001	*0.19*	9.34	2137	<0.001	*0.47*	4.30	1809	<0.001	*0.21*
Native language	44.22	1	<0.001	*0.18*	66.53	1	<0.001	*0.18*	0.55	1	0.459	0.02
Educational attainment	7.70	3	0.053	0.06	55.41	3	<0.001	*0.16*	41.35	3	<0.001	*0.15*
No vs. low	0.06	1	0.809	0.01	37.32	1	<0.001	*0.24*	33.65	1	<0.001	*0.26*
No vs. medium	2.82	1	0.093	0.05	47.50	1	<0.001	*0.19*	23.92	1	<0.001	*0.14*
No vs. high	0.16	1	0.686	0.02	44.21	1	<0.001	*0.31*[a]	31.02	1	<0.001	*0.29*
Low vs. medium	5.76	1	0.016	0.07	0.04	1	0.848	0.01	4.81	1	.028	0.06
Low vs. high	0.57	1	0.449	0.04	1.64	1	.201	0.05	0.05	1	.820	0.01
Medium vs. high	0.85	1	0.358	0.03	1.60	1	.206	0.03	4.65	1	.031	0.06
Employment status	21.25	2	<0.001	*0.11*	14.72	2	<0.001	0.08	56.32	2	<0.001	*0.18*
Employed vs. unemployed	12.43	1	<0.001	*0.10*	0.13	1	0.716	0.01	15.24	1	<0.001	*0.11*
Employed vs. other	3.64	1	0.056	0.05	12.61	1	<0.001	0.09	27.59	1	<0.001	*0.14*
Unemployed vs. other	20.49	1	<0.001	*0.17*	9.93	1	0.002	*0.11*	52.14	1	<0.001	*0.28*
Occupational status	2.37	1856	0.018	− 0.11	2.27	2137	0.023	− 0.10	0.34	1809	.735	0.02

t-value for age and occupational status, X^2 for gender, native language, educational attainment, and employment status; φ for gender, native language, educational attainment, and employment status; *ES* Effect size (Cramer's *V* for educational attainment and employment status; φ for gender, native language, educational attainment, and employment status; Cohen's *d* for age and occupational status); Variables highlighted in italic are significant in terms of small effect sizes, variables highlighted in italic with [a] are significant in terms of medium effect sizes

employment. Whereas the proportion of unemployed adults was almost equally high in LEO (17.17%) and NEPS (19.36%), it was about seven to nine percentage points lower in PIAAC (10.29%), although effect sizes were small. Furthermore, with 29.97% in PIAAC, the proportion of low-literate adults in another employment status was slightly higher compared to LEO (with a difference of 6.02%) and NEPS (with a difference of 12.62%), although again effect sizes were small.

The association between low literacy and relevant correlates

Figure 2 shows the AMEs from our logistic regressions among all adults (Model 1) and among employed adults only (Model 2) (also see Tables 7, 8 in the Appendix).

First, we found low literacy to be related to age. Across all three surveys, we found that the risk of being in the low-literacy group increased with age; in all studies, it was higher in the oldest age group (54 to 64 years; 8 to 12%) than in the youngest age group (24 to 33 years, 2 to 3%). Second, the results show that male respondents of LEO and NEPS were more likely to be low-literate than female respondents; however, the differences are relatively small with male respondents being 4 to 7% more likely to belong to the low-literacy group than female respondents. There are no gender differences for PIAAC. Among the sociodemographic factors, we found that for LEO and PIAAC, native language had the largest association with low literacy. For adults with a non-German background, the probability of being in the low-literacy group was almost twice as high for respondents of PIAAC and LEO as for NEPS respondents. However, the association between native language and low literacy, with non-German respondents being 9% more likely to belong to the low-literacy group than native respondents, in NEPS was still relatively high compared to age group 2 (34 to 43 years) and age group 3 (44 to 53 years) with 3% and 4% respectively, and compared to gender, with men being 4% more likely to belong to the low-literacy group.

Compared to sociodemographic factors, the association between most socioeconomic factors and low literacy was larger and showed similar patterns across the surveys. Results regarding the educational attainment reflect previous findings according to which the probability of belonging to the low-literacy group increased with lower educational attainment. As outlined, our reference group were adults with a high educational attainment. Low-literate respondents with no educational attainment were between 44 to 60% more likely to be low-literate compared to our reference group. Adults with a low educational attainment were on average 21 to 32% more likely to belong to the low-literacy group than adults with a high level of educational attainment. In other words, the probability of being low-literate was four to six times higher for adults with no educational degree and two to three times higher for those with a low degree than for those with a medium educational degree. With respect to the employment status, the probability of being in the low-literacy group was 10 to 15% higher for unemployed adults than for employed adults.

Model 2 additionally considered the occupational status. Whereas a low occupational status was linked to a considerably higher risk of having low literacy proficiency, introducing occupational status and restricting the sample to employed respondents also slightly reduced the magnitude of almost all other variables compared to Model 1. The AME for adults with a medium educational attainment shrunk the most for PIAAC with

from 27 to 9%, but the effect remained comparably high for adults with no educational attainment. As expected from previous literature, adults with a low occupational status were 10 to 15% more likely to belong to the low-literacy group compared to adults with the highest occupational status. The association between a medium and high occupational status was relatively smaller, ranging 3 to 5% for medium vs. very high and from 2 to 3% for high vs. very high.

Discussion

Given the persistently high number of low-literate adults in Germany compared to other Western OECD countries (Durda et al. 2020; Grotlüschen et al. 2019a; Grotlüschen and Riekmann 2011; OECD 2013b), it is of high relevance to researchers, policymakers and practitioners to use comparable databases to understand the phenomenon of low literacy, for example, in terms of risk and protective factors for the development of low literacy. The three large-scale surveys available in Germany for this purpose–LEO (Grotlüschen and Riekmann 2011), PIAAC (OECD 2013b), and NEPS (Blossfeld and oßbach, 2019)–offer various and complementary potentials for analyses, but the implications derived from those surveys decisively depend on the comparability of the low-literacy samples. Therefore, our article aimed to garner new insights about the comparability of low-literate adults across LEO, PIAAC, and NEPS.

Synthesis of the findings

The most important finding of our study is that – despite divergent sample representations and measurement approaches of the three surveys–the group of low-literate adults was highly similar across LEO, PIAAC, and NEPS: Data from all three surveys indicated that low-literate adults were more likely to have an older age and non-native language background, lower levels of educational attainment and occupational status and higher rates of unemployment than literate adults. At the same time, the results indicated that low-literate adults are a heterogeneous group that can be found in all sociodemographic and socioeconomic groups.

Our results also revealed small differences between the surveys. First, we found that gender differences within the low-literacy group were survey-specific. Whereas low literacy was higher among males than females in LEO and NEPS, gender differences were hardly apparent in PIAAC. In line with the interpretation by Solheim and Lundetræ (2018), these gender differences might be an artefact of test construction. Because the reading items of PIAAC have a higher proportion of non-continuous texts and digital texts that are assumed to be more male-friendly, we probably did not find gender differences for low-literate adults in PIAAC but a higher proportion of male low-literate adults in LEO and NEPS. Second, we found more pronounced differences between native and non-native speakers in LEO compared to PIAAC and NEPS. As with the gender differences, the differences in the assessments give reasons to assume that the magnitude of the observed differences between native and non-native speakers appear to be associated with certain assessment features. LEO includes a high proportion of items that require respondents to document their writing skills. This might make the LEO test more native-friendly because of the added difficulty of writing in a foreign language

(Grotlüschen et al. 2012). PIAAC and NEPS, by contrast, measure reading literacy, and do not require respondents to do any writing, which might make both tests more non-native-friendly. Finally, most striking were the differences in the distribution of the educational attainment among the low-literacy samples, with a smaller proportion of adults with no educational attainment and a relatively higher proportion of adults with a high educational attainment in NEPS compared to PIAAC and LEO. It can be assumed that these deviations arose from a considerably lower response rate of adults with no educational attainment in the full sample of NEPS compared to LEO and PIAAC.

Further, we found that the patterns of low literacy and its correlates were predominantly robust across LEO, PIAAC, and NEPS. This is a good and encouraging result, again speaking in favor of the comparability of the subgroup of adults with low literacy. The convergence of the main findings between the surveys can be taken as quite compelling, given that the surveys use different assessments and were designed for different purposes. It comes as no surprise that no educational attainment stood out as the most important correlate of low literacy, followed by a low educational attainment. With some variation across the three surveys, the other most important correlates of low-literacy were a non-German language background, unemployment and low occupational status, and higher age. In contrast, the correlates that demonstrated the least important correlates of low literacy were gender, other employment status and younger age.

Implications for educational research and policy debate

What do these results indicate for use in educational research, policy and monitoring? Until now, it has been largely unclear how the group of adults with low literacy in LEO is comparable to adults with low literacy in PIAAC and NEPS. To the best of our knowledge, no previous study has used the three data sources for comparative analyses, leaving a large potential untapped. In the present study, we shed light on this research gap, and showed that overall, the subgroup of adults with low literacy is comparable across LEO, PIAAC, and NEPS; therefore, all surveys can be used for more in-depth investigations within this group, thereby unfolding the full complementary value of all three studies. However, the results have also revealed limitations in potential uses. For example, if the focus is specifically on the group of adults with no educational attainment, then LEO and PIAAC should be used because of the limited number of cases in NEPS. If, on the other hand, the focus is on reading literacy, then PIAAC and NEPS provide more valid information than LEO due to its focus on reading and writing.

Moreover, our results contribute to the debate on the magnitude of gender differences and differences regarding the proportion of non-native speakers within the low-literacy group and why it might differ across surveys. According to our findings, the assessment of literacy is a moderating factor that may help to explain heterogeneous findings in previous research (e.g., Grotlüschen and Riekmann 2012; OECD 2013b; Thums et al. 2020). The most important implication of our findings consists of creating transparency about the assessment of (low) literacy and its impact on proficiency differences when using one of these surveys. Different assessments reflect different parts of the respondents' literacy potential. Therefore the particular strengths of each survey should be used for following different research questions in order to avoid invalid conclusions that could arise if the reported differences in literacy are taken out of context.

Among the correlates of low literacy, several factors come to the fore that offer starting points for the prevention and promotion of low literacy. These include formal educational institutions, informal and non-formal learning opportunities provided by the workplaces or employment agencies. For some of the adults with non-German language background, opportunities to learn German may help minimize disadvantages in literacy. Moreover, the fact that gender differences were small, if any, when controlling for social and educational background implies that any correlation between gender and literacy proficiency should always be interpreted with this contextual information in mind. In other words, social and educational background matter for interpreting the higher representation of men within the group of low-literate adults.

Limitations of the present study and future research

The present study has several limitations. In this study, we have argued that a range of characteristics at the level of measurement cause variations in the low-literacy samples. Unfortunately, we were not able to empirically test this assumption, since not all surveys are linked on a common scale. However, this would be highly interesting for future research, because a joint empirical examination of the assessments not only allows statements to be made on construct equivalence but differences in the magnitude of low literacy can also be quantified more precisely against the background of identical scaling. Another limitation concerns the role of accuracy in the proficiency level assignment. Several factors can influence a person's proficiency level assignment, such as the range and the number of the test items, the choice of the standard-setting procedure, or the ability distribution to name a few (e.g., Cizek 2012; Wu and Nguyen 2019). We have pointed out some of these differences. For example, PIAAC and LEO use plausible values to account for measurement uncertainty, NEPS has examined the accuracy in the proficiency level assignment by means of a validation study. However, further research is needed on the extent to which the uncertainty in the proficiency level assignment explains the differences between those surveys. Third, this study only considered a small number of indicators for low reading proficiency, as only those limited set of indicators were measured similarly across the surveys. For further research, other factors, such as reading practices or reading-related skills (such as vocabulary), should be considered too. And finally, this study is not based on longitudinal data, so no conclusion with regard to causality and life course mechanism about the development of low literacy can be drawn.

Conclusions

Our study demonstrates that the group of low-literate adults in the three German large-scale surveys of LEO, PIAAC, and NEPS are largely comparable in terms of their sociodemographic and socioeconomic characteristics. Differences in the composition of these groups across the three surveys were mostly small and can likely be traced back to differences in the sample representation and measurement of (low) literacy. These differences should be considered when using the surveys for comparative research and policy purposes. Nevertheless, our study showed that we do not

compare apples with oranges when dealing with low-literate adults across the three surveys.

Author details
[1] LIfBi–Leibniz Institute for Educational Trajectories, Bamberg, Germany. [2] GESIS–Leibniz Institute for the Social Sciences, Mannheim, Germany. [3] University of Hamburg, Hamburg, Germany.

Appendix

See (Tables 4, 5, 6, 7, 8)

Table 4 Overview of variables and how they are measured

Variable	Survey instrument
Low literacy	
LEO	Alpha-level 3 and below (Hartig and Riekmann 2012), 5 variables derived from 5 plausible values for literacy skills for each respondent
PIAAC	Proficiency level 1 and below (Yamamoto et al. 2012b), 10 variables derived from 10 plausible values for literacy skills for each respondent
NEPS	Proficiency level 1 (Durda et al. 2020), 1 variable derived from 1 WLE value for literacy skill for each respondent
Gender	
LEO	Ascertained by information on the gender
PIAAC	Recoded into 0: female, 1: male
NEPS	
Age	
LEO	Ascertained by information on the date of birth and the interview date For logistic regression recoded into: 1: 24 to 33 years, 2: 34 to 43 years, 3: 44 to 53 years, 4: 54 to 64 years
PIAAC	
NEPS	
Native language	
LEO	Which language did you learn first in your childhood? (1: German, 2: another language, 3: bilingual with German and other language, 4: bilingual, without German, 9: not specified) Recoded into 0: German, 1: non-German
PIAAC	What is your mother tongue? That is, the first language you spoke at home during your childhood and that you still understand (1: German, 2: Turkish, 3: Italian, 4: Polish, 5: Greek, 6: Serbian, 7: Croatian, 8: Russian, 9: another language); Recoded into 0: German, 1: non-German
NEPS	What language did you learn as a child in your family? Recoded into 0: German, 1: non-German
Highest educational attainment	
LEO	International Standard Classification of Education (ISCED-97)
PIAAC	Recoded into 3: no (ISCED 1 and below), 2: low (ISCED 2), 1: medium (ISCED 3/4), 0: high (ISCED 5/6)
NEPS	
Occupational status	
LEO	International Socioeconomic Index of Occupational Status (ISEI-08, Ganzeboom 2010)
PIAAC	
NEPS	For logistic regression recoded into: 4: low occupational status (11.56 to 30.91), 3: medium occupational status (30.92 to 50.25), 2: high occupational status (50.26 to 69.60) and 1: very high occupational status (69.61 to 88.96). Note: In NEPS, determined from the spell type (sptype) in the episode split of the biography data (for details, see Rompczyk and Kleinert 2017), in cases of multiple spells, employment spell with the longest duration was selected
Employment status	
LEO	Are you personally at present … (1: full-time employed, 2: part-time employed, 3: unemployed, 4: on parental leave, 5: trainee in company apprenticeship, … 11: other, 99: not specified) Recoded into 1: employed, 2: unemployed, 3: other
PIAAC	Please have a look at this list and tell me which statement best describes your current situation. Decide which statement you think best suits you (1: full-time employed, 2: part-time working: 3: unemployed, 3: school pupil, student, …10: other) Recoded into 1: employed, 2: unemployed, 3: other
NEPS	Employment status is determined from the spell type (sptype) in the episode split of the biography data (for details, see Rompczyk and Kleinert 2017) Recoded into 1: employed, 2: unemployed, 3: other
Sampling weights	
LEO	Nonresponse weights [pgew]
PIAAC	Nonresponse weights [SPFWT0] and replicate weights [SPFWT1-SPFWT80]
NEPS	Nonresponse weights [w_t3_cal or w_t5_cal]

Table 4 (continued)

The international master version of the PIAAC background questionnaire can be accessed at: https://www.oecd.org/skill s/piaac/BQ_MASTER.HTM. The NEPS questionnaire can be accessed at: https://www.neps-data.de/Data-Center/Data-and-Documentation/Starting-Cohort-Adults/Documentation. The LEO questionnaire can be accessed at: https://leo.blogs .uni-hamburg.de/?attachment_id=1081

Table 5 Interpreting of Effect Sizes for Phi *φ*, Cramer's *V*, Cohens's *f*, and Cohen's *d*

Effect sizes	Phi *φ*	Cramer's *V*	Cohen's *f*	Cohen's d
Small	$\varphi = 0.10$	df $= 1$ ($V = 0.10$) df $= 2$ ($V = 0.07$) df $= 3$ ($V = 0.06$) df $= 4$ ($V = 0.05$) df $= 5$ ($V = 0.04$)	$f = 0.10$	$|d| = 0.2$
Medium	$\varphi = 0.30$	df $= 1$ ($V = 0.30$) df $= 2$ ($V = 0.21$) df $= 3$ ($V = 0.17$) df $= 4$ ($V = 0.05$) df $= 4$ ($V = 0.15$) df $= 5$ ($V = 0.13$)	$f = 0.25$	$|d| = 0.5$
Large	$\varphi = 0.50$	df $= 1$ ($V = 0.50$) df $= 2$ ($V = 0.35$) df $= 3$ ($V = 0.29$) df $= 4$ ($V = 0.25$) df $= 5$ ($V = 0.22$)	$f = 0.40$	$|d| = 0.8$

df for Cramer's *V* it is referred to as DF* $= (R-1)$ or $(C-1)$, whereas the df for Chi-square is defined as DF $= (R-1) + (C-1)$ (Cohen 1988)

Table 6 Logistic Regression Predicting Low Literacy for LEO, PIAAC, and NEPS among all Adults

Effect	AME	SE	95%		p
			LL	UL	
LEO					
Age 34–43 (Ref. 24–33)	0.03	0.01	0.01	0.05	0.005
Age 44–53 (Ref. 24–33)	0.07	0.01	0.05	0.09	< 0.001
Age 54–64 (Ref. 24–33)	0.08	0.01	0.06	0.11	< 0.001
Male	0.07	0.01	0.05	0.09	< 0.001
Non-German	0.17	0.02	0.13	0.21	< 0.001
Medium educational attainment (Ref. high)	0.08	0.01	0.06	0.10	< 0.001
Low educational attainment (Ref. high)	0.22	0.02	0.17	0.27	< 0.001
No educational attainment (Ref. high)	0.50	0.04	0.41	0.58	< 0.001
Unemployed (Ref. employed)	0.11	0.02	0.07	0.16	< 0.001
Other employment status (Ref. employed)	0.04	0.01	0.01	0.06	0.001
PIAAC					
Age 34–43 (Ref. 24–33)	0.02	0.02	− 0.02	0.06	0.270
Age 44–53 (Ref. 24–33)	0.10	0.02	0.05	0.14	< 0.001
Age 54–64 (Ref. 24–33)	0.12	0.02	0.07	0.17	< 0.001
Male	0.01	0.02	− 0.02	0.05	0.415
Non-German	0.18	0.03	0.12	0.24	< 0.001
Medium educational attainment (Ref. high)	0.13	0.01	0.11	0.16	< 0.001
Low educational attainment (Ref. high)	0.32	0.04	0.23	0.41	< 0.001
No educational attainment (Ref. high)	0.60	0.07	0.46	0.74	< 00.001
Unemployed (Ref. employed)	0.10	0.03	0.03	0.17	0.005
Other employment status (Ref. employed)	0.04	0.02	0.00	0.09	0.037
NEPS					
Age 34–43 (Ref. 24–33)	0.03	0.01	0.01	0.04	0.007
Age 44–53 (Ref. 24–33)	0.04	0.01	0.02	0.05	< 0.001
Age 54–64 (Ref. 24–33)	0.10	0.01	0.08	0.12	< 0.001
Male	0.04	0.01	0.03	0.05	< 0.001
Non-German	0.09	0.01	0.07	0.11	< 0.001
Medium educational attainment (Ref. high)	0.10	0.01	0.08	0.11	< 0.001
Low educational attainment (Ref. high)	0.21	0.02	0.16	0.25	< 0.001
No educational attainment (Ref. high)	0.44	0.07	0.31	0.58	< 0.001
Unemployed (Ref. employed)	0.15	0.02	0.12	0.19	< 0.001
Other employment status (Ref. employed)	0.05	0.01	0.03	0.07	< 0.001

AME Average Marginal Effect; LL and UL indicate the lower and upper limits of a confidence interval, respectively

Table 7 Logistic Regression Predicting Low Literacy for LEO, PIAAC, and NEPS among Adults in Employment

Effect	AME	SE	95%		p
			LL	UL	
LEO					
Age 34–43 (Ref. 24–33)	0.02	0.02	− 0.02	0.05	0.309
Age 44–53 (Ref. 24–33)	0.04	0.01	0.02	0.06	0.001
Age 54–64 (Ref. 24–33)	0.04	0.02	0.01	0.08	0.019
Male	0.06	0.01	0.04	0.07	< 0.001
Non-German	0.12	0.02	0.09	0.16	< 0.001
Medium educational attainment (Ref. high)	0.04	0.02	0.00	0.07	0.032
Low educational attainment (Ref. high)	0.17	0.04	0.09	0.24	< 0.001
No educational attainment (Ref. high)	0.39	0.06	0.25	0.52	< 0.001
High occupational status (Ref. very high)	0.03	0.02	− 0.01	0.07	0.103
Medium occupational status (Ref. very high)	0.05	0.02	0.01	0.09	0.028
Low occupational status (Ref. very high)	0.10	0.02	0.07	0.14	< 0.001
PIAAC					
Age 34–43 (Ref. 24–33)	− 0.01	0.02	− 0.05	0.04	0.759
Age 44–53 (Ref. 24–33)	0.06	0.02	0.02	0.11	0.010
Age 54–64 (Ref. 24–33)	0.08	0.03	0.01	0.15	0.036
Male	0.00	0.02	− 0.04	0.03	0.880
Non-German	0.17	0.04	0.08	0.26	0.001
Medium educational attainment (Ref. high)	0.07	0.02	0.03	0.11	< 0.001
Low educational attainment (Ref. high)	0.18	0.05	0.09	0.27	< 0.001
No educational attainment (Ref. high)	0.43	0.10	0.23	0.64	0.00
High occupational status (Ref. very high)	0.03	0.03	− 0.02	0.08	0.266
Medium occupational status (Ref. very high)	0.05	0.03	− 0.02	0.12	0.152
Low occupational status (Ref. very high)	0.15	0.03	0.08	0.21	< 0.001
NEPS					
Age 34–43 (Ref. 24–33)	0.02	0.01	0.00	0.03	0.053
Age 44–53 (Ref. 24–33)	0.03	0.01	0.01	0.05	< 0.001
Age 54–64 (Ref. 24–33)	0.08	0.01	0.07	0.10	< 0.001
Male	0.03	0.01	0.02	0.05	< 0.001
Non-German	0.08	0.01	0.06	0.10	< 0.001
Medium educational attainment (Ref. high)	0.04	0.01	0.03	0.06	< 0.001
Low educational attainment (Ref. high)	0.11	0.02	0.07	0.15	< 0.001
No educational attainment (Ref. high)	0.33	0.09	0.16	0.50	< 0.001
High occupational status (Ref. very high)	0.02	0.01	0.01	0.03	0.006
Medium occupational status (Ref. very high)	0.03	0.01	0.01	0.04	< 0.001
Low occupational status (Ref. very high)	0.13	0.01	0.11	0.15	< 0.001

AME Average Marginal Effect; LL and UL indicate the lower and upper limits of a confidence interval, respectively

Table 8 Proportion of missing values for analytic samples of LEO, PIAAC, and NEPS in %

	LEO	PIAAC	NEPS
Literacy	0.00	0.00	0.00
Gender	0.00	0.00	0.01
Age	0.00	0.00	0.00
Native language	0.00	1.84	0.06
Educational attainment	0.36	1.89	0.34
Employment status	0.00	1.75	1.40
Occupational status	1.20	0.70	3.22

Proportion of missing values for occupational status refers to respondents in employment

References

Arendt, J. N., Rosholm, M., & Jensen, T. P. (2008). The importance of literacy for employment and unemploysment duration. *Nationaløkonomisk tidsskrift, 146*(1), 22–26.

Arrow, K. J. (1973). Higher education as a filter. *Journal of Public Economics, 2*(3), 193–216.

Authoring Group Educational Reporting. (2016). *Education in Germany 2016*. Bertelsmann: An indicator-based report including an analysis on education and migration.

Barone, C., & Van de Werfhorst, H. G. (2011). Education, cognitive skills and earnings in comparative perspective. *International Sociology, 26*(4), 483–502. https://doi.org/10.1177/0268580910393045

Barzillai, M., Thomson, J. M., Schroeder, S., & Broek, P. W. van den (Eds.). (2018). *Learning to read in a digital world.* John Benjamins Publishing Company.

Bilger, F., Jäckle, R., von Rosenbladt, B., & Strauß, A. (2012). Studiendesign, Durchführung und Methodik der leo.–Level-One-Studie. In A. Grötlüschen, & W. Riekmann (Eds.), *Funktionaler Analphabetismus in Deutschland: Ergebnisse der ersten leo - Level-One Studie.* (pp. 77–105). Waxmann.

Blömeke, S., & Gustafsson, J.-E. (Eds.). (2017). *Standard setting in education: The Nordic countries in an international perspective.* Springer.

Blossfeld, G. J., Blossfeld, P. N., & Blossfeld, H.-P. (2019). A sociological perspective on education as a lifelong process. In R. Becker (Ed.), *Research handbook on the sociology of education* (pp. 18–34). Cheltenham: Edward Elgar Publishing.

Blossfeld, H.-P., & Roßbach, H.-G. (2019). *Education as a lifelong process: The German National Educational Panel Study (NEPS)*. VS Verlag für Sozialwissenschaften.

Braun, H., & von Davier, M. (2017). The use of test scores from large-scale assessment surveys: psychometric and statistical considerations. *Large-Scale Assessments in Education, 5*(1), 17. https://doi.org/10.1186/s40536-017-0050-x

Buchmann, C., DiPrete, T. A., & McDaniel, A. (2008). Gender inequalities in education. *Annual Review of Sociology, 34*(1), 319–337. https://doi.org/10.1146/annurev.soc.34.040507.134719

Buddeberg, K., Dutz, G., Grotlüschen, A., Hartig, J., Heilmann, L., Stammer, C., (2020). Verhältnis der Kompetenzstufen in PIAAC und LEO (2018). In A. Grotlüschen & K. Buddeberg (Eds.), *LEO 2018 – Leben mit geringer Literalität (pp 353–367)*. wbv: Bielefeld.

Bus, A. G., & van IJzendoorn, M. H., & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research, 65*(1), 1–21. https://doi.org/10.3102/00346543065001001

Bynner, J., & Parsons, S. (1998). *Use it or lose it? The impact of time out of work on literacy and numeracy skills.* Basic Skills Agency.

Bynner, J., & Parsons, S. (2006). *New light on literacy and numeracy*. National Research and Development Centre for adult literacy and numeracy: Summary report.

Calero, J., & Choi, Á. (2017). The distribution of skills among the European adult population and unemployment: A comparative approach. *European Journal of Education, 52*(3), 348–364. https://doi.org/10.1111/ejed.12222

Carstensen, C. H., Gaasch, J.-C., & Rothaug, E. (2017). *Kompetenzmodellierung in PIAAC-L aktueller Stand Welle 3 (2016)*. Unpublished internal document.

Cheema, J. R. (2014). Some general guidelines for choosing missing data handling methods in educational research. *Journal of Modern Applied Statistical Methods, 13*(2), 53–75. https://doi.org/10.22237/jmasm/1414814520

Choi, H., & Feng, J. (2015). General Slowing Hypothesis. In S. K. Whitbourne (Ed.), *The Encyclopedia of Adulthood and Aging* (pp 1–4). Wiley. https://doi.org/10.1002/9781118521373.wbeaa198

Cizek, G. J. (2012). *Setting performance standards: Foundations, methods, and innovations* (2nd edition). Routledge. https://doi.org/10.4324/9780203848203

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edition). Lawrence Erlbaum.

Cortina, K. S., Baumert, J., Leschinsky, A., Mayer, K. U., & Trommer, L. (Eds.). (2005). *Das Bildungswesen in der Bundesrepublik Deutschland: Strukturen und Entwicklungen im Überblick* (2nd edition). Rowohlt-Taschenbuch-Verlag.

Desjardins, R., & Warnke, A. J. (2012). *Ageing and skills: A review and analysis of skill gain and skill loss over the lifespan and over time* (OECD Education Working Papers No. 72). https://doi.org/10.1787/5k9csvw87ckh-en

Ditton, H., & Krüsken, J. (2010). Bildungslaufbahnen im differenzierten Schulsystem – Entwicklungsverläufe von Laufbahnempfehlungen und Bildungsaspirationen in der Grundschulzeit. In J. Baumert, K. Maaz, & U. Trautwein (Eds.), *Bildungsentscheidungen* (pp. 74–102). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92216-4_4

Durda, T., Artelt, C., Lechner, C., Rammstedt, B., & Wicht, A. (2020). Proficiency level descriptors for low reading proficiency: An integrative process model. *International Review of Education*. https://doi.org/10.1007/s11159-020-09834-1

Flisi, S., Goglio, V., Meroni, E. C., & Vera-Toscano, E. (2019). Cohort patterns in adult literacy skills: How are new generations doing? *Journal of Policy Modeling, 41*(1), 52–65. https://doi.org/10.1016/j.jpolmod.2018.10.002

Frijters, J. C., Brown, E., & Greenberg, D. (2019). Gender differences in the reading motivation of adults with low literacy Sskills. In: D. Perin (Ed.), *The wiley handbook of adult literacy* (pp. 63–87). Hoboken: Wiley. Doi: https://doi.org/10.1002/9781119261407.ch3

Ganzeboom, H. B. G. (May, 2010). *A new international socioeconomic index [ISEI] of occupational status for the international standard classification of occupation 2008 [ISCO-08] constructed with data from the ISSP 2002–2007; With an analysis of quality of occupational measurement in ISSP.* Conference of International Social Survey Programme, Lisbon.

Gauly, B., & Lechner, C. M. (2019). Self-perfection or self-selection? Unraveling the relationship between job-related training and adults' literacy skills. *PLoS ONE.* https://doi.org/10.1371/journal.pone.0215971

Gehrer, K., Zimmermann, S., Artelt, C., & Weinert, S. (2013). NEPS framework for assessing reading competence and results from an adult pilot study. *Journal for Educational Research Online, 5*(2), 50–79.

Graham, J. W., Cumsille, P. E., Shevock, A. E. (2012). Methods for handling missing data. In I. Weiner (Ed), *Handbook of psychology* (2nd edition). Hoboken: Wiley. https://doi.org/10.1002/9781118133880.hop202004

Green, F. (2013). *Skills and skilled work: an economic and social analysis.* Oxford: Oxford University Press.

Grotlüschen, A., Buddeberg, K., Dutz, G., Heilmann, L., & Stammer, C. (2019a). *LEO: Living with low literacy.* https://leo.blogs.uni-hamburg.de/wp-content/uploads/2019/07/LEO_2018_Living_with_Low_Literacy.pdf

Grotlüschen, A, Buddeberg, K., Dutz, G., Heilmann, L. M., & Stammer, C. (2019b). Practices and Competencies–evidence from an adult literacy survey in Germany. In M. Schemmann (Ed.), *Internationales Jahrbuch der Erwachsenenbildung / International year book of adult education* (pp. 17–34). wbv.

Grotlüschen, A., Buddeberg, K., Dutz, G., Heilmann, L. M., & Stammer, C. (2020). Low literacy in Germany Results from the second German literacy survey. *European Journal for Research on the Education and Learning of Adults.* https://doi.org/10.3384/rela.2000-7426.rela9147

Grotlüschen, A., & Riekmann, W. (2011). *Leo.–Level-one study. Literacy of adults at the lower rungs of the ladder.* University of Hamburg. https://leo.blogs.uni-hamburg.de/wp-content/uploads/2011/12/leo-Press-brochure15-12-2011.pdf

Grotlüschen, A., & Riekmann, W. (Eds.). (2012). *Funktionaler Analphabetismus in Deutschland: Ergebnisse der ersten leo.–Level-One Studie.* Waxmann.

Grotlüschen, A., Mallows, D., Reder, S., & Sabatini, J. (2016). *Adults with low proficiency in literacy or numeracy* (OECD Education Working Papers No. 131). https://doi.org/10.1787/5jm0v44bnmnx-en.

Grotlüschen, A., Riekmann, W., & Buddeberg, K. (2012). Leo.–Level-One Studie: Methodische Herausforderungen. In A. Grotlüschen, & W. Riekmann (Eds.), *Funktionaler Analphabetismus in Deutschland: Ergebnisse der ersten leo.–Level-One Studie* (pp. 54–76). Waxmann.

Groves, R. M., & Lyberg, L. (2010). Total survey error: Past, present, and future. *Public Opinion Quarterly, 74*(5), 849–879. https://doi.org/10.1093/poq/nfq065

Haberkorn, K., Pohl, S., Hardt, K., & Wiegand, E. (2012). *NEPS technical report for reading: Scaling results of Starting Cohort 4 in ninth grade* (NEPS Working Paper No. 16). Leibniz Institute for Educational Trajectories, National Educational Panel Study.

Hammon, A., Zinn, S., Aßmann, C., & Würbach, A. (2016). *Samples, weights, and nonresponse: The Adult Cohort of the National Educational Panel Study (Wave 2 to 6)* (NEPS Survey Paper No. 7). Leibniz Insitute for Educational Trajectories, National Educational Panel Study.

Hartig, J., & Riekmann, W. (2012). Bestimmung der Level-Grenzen in der leo. – Level-One Studie. In A. Grotlüschen, & W. Riekmann (Eds.), *Funktionaler Analphabetismus in Deutschland: Ergebnisse der ersten leo. - Level-One Studie* (pp. 106–121). Waxmann.

Karantonis, A., & Sireci, S. G. (2006). The Bookmark standard-setting method: A literature review. *Educational Measurement: Issues and Practice, 25*(1), 4–12. https://doi.org/10.1111/j.1745-3992.2006.00047.x

Kirsch, I. S., Yamamoto, K., & Garber, D. (2013). Chapter 1: PIAAC assessment design. In OECD (Ed.), *Technical report of the survey of adult skills (PIAAC)* (2nd edition, p. Chapter 1). OECD.

Koller, I., Haberkorn, K., & Rohm, T. (2014). *NEPS technical report for reading: Scaling results of Starting Cohort 6 for Adults in main study 2012* (NEPS Working Paper No. 48). Leibniz Institute for Educational Trajectories, National Educational Panel Study.

Lenters, K. (2006). Resistance, struggle, and the adolescent reader. *Journal of Adolescent & Adult Literacy, 50*(2), 136–146. https://doi.org/10.1598/JAAL.50.2.6

Lin, Y. (2006). The Bookmark procedure for setting cut-scores and finalizing performance standards: Strengths and weaknesses. *The Alberta Journal of Educational Research, 52*(1), 39–52.

Martin, S., Lechner, C. M., Kleinert, C., & Rammstedt, B. (2020). Literacy skills predict probability of refusal in follow-up wave: Evidence from two longitudinal assessment surveys. *International Journal of Social Research Methodology.* https://doi.org/10.1080/13645579.2020.1763715

McIntosh, S. (2001). Measuring and assessing the impact of basic skills on labour market outcomes. *Oxford Economic Papers, 53*(3), 453–481. https://doi.org/10.1093/oep/53.3.453

Mohadjer, L., Krenzke, T., & Van de Kerckhove, W. (2013a). Chapter 16: Indicators of the quality of the sample data. In OECD (Ed.), *Technical report of the survey of adult skills* (PIAAC) (2nd edition, p. Chapter 16). OECD.

Mohadjer, L., Krenzke, T., Van de Kerckhove, W., & Hsu, W. (2013b). Chapter 15: Survey weighting and variance estimation. In OECD (Ed.), *Technical report of the survey of adult skills (PIAAC)* (2nd edition, p. Chapter 15). OECD.

Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin, 137*(2), 267–296. https://doi.org/10.1037/a0021890

Mullis, I. V. S., Martin, M. O., Foy, P., & Hopper, M. (2017). *PIRLS 2016 international results in reading.* https://timssandpirls.bc.edu/pirls2016/international-results/

OECD (2013a). *Technical report of the survey of adult skills (PIAAC). OECD skills studies*. OECD. https://www.oecd.org/skills/piaac/PIAAC_Technical_Report_2nd_Edition_Full_Report.pdf.

OECD. (2013a). OECD skills outlook 2013: First results from the survey of adult skills. *OECD*. https://doi.org/10.1787/9789264204256-en

OECD. (2013b). The survey of adult skills: Reader's companion. *OECD*. https://doi.org/10.1787/9789264204027-en

OECD. (2019a). Skills matter: Additional results from the survey of adult skills. *OECD*. https://doi.org/10.1787/1f029d8f-en

OECD. (2019b). PISA 2018 results: What students know and can do. *OECD*. https://doi.org/10.1787/5f07c754-en

Paccagnella, M. (2016). *Age, ageing and skills: Results from the survey of adult skills* (OECD Education Working Papers No. 132). OECD. https://doi.org/10.1787/5jm0q1n38lvc-en

Perry, A., & Gauly, B. (2019). *Analysing poor reading skills: a comparison between PIAAC literacy levels and reading components*. Survey Methods: Insights from the Field. https://doi.org/10.13094/SMIF-2019-00020

Pfost, M., & Artelt, C. (2014). Reading literacy development in secondary school and the effect of differential institutional learning environments. In M. Pfost, C. Artelt, & S. Weinert (Eds.), *The development of reading literacy from early childhood to adolescence. Empirical findings from the Bamberg BiKS longitudinal study* (pp. 229–277). University of Bamberg Press.

PIAAC Literacy Expert Group. (2009). *PIAAC literacy: A conceptual framework* (OECD Education Working Papers No. 34). OECD. https://www.oecd-ilibrary.org/education/piaac-literacy-a-conceptual-framework_220348414075

Pohl, S., & Carstensen, C. H. (2012). *NEPS technical report: Scaling the data of the competence tests* (NEPS Technical Report No. 14). Leibniz Institute for Educational Trajectories, National Educational Panel Study.

Pohl, S., Gräfe, L., & Rose, N. (2014). Dealing with omitted and not-reached items in competence tests: Evaluating approaches accounting for missing responses in item response theory models. *Educational and Psychological Measurement, 74*(3), 423–452. https://doi.org/10.1177/0013164413504926

Rauch, D. P., & Hartig, J. (2010). Multiple-choice versus open-ended response formats of reading test items: A two-dimensional IRT analysis. *Psychological Test and Assessment Modeling, 52*(4), 354–379.

Reder, S. M. (2009). The development of literacy and numeracy in adult life. In S. M. Reder, & J. M. Bynner (Eds.), *Tracking adult literacy and numeracy skills: Findings from longitudinal research.* (pp. 59–84). Routledge.

Reder, S., & Gauly., B., & Lechner, C. M. . (2020). Practice makes perfect: Practice engagement theory and the development of adult literacy and numeracy proficiency. Manuscript accepted for publication. *International Review of Education*. https://doi.org/10.1007/s11159-020-09830-5

Robitzsch, A. (2011). Zu nichtignorierbaren Konsequenzen des (partiellen) Ignorierens fehlender Item Responses im Large-Scale Assessment. In B. Suchań, C. Wallner-Paschon, & C. Schreiner (Eds.). *PIRLS & TIMSS 2011—Die Kompetenzen in Lesen, Mathematik und Naturwissenschaft am Ende der Volksschule: Österreichischer Expertenbericht* (pp. 55–64). Leykam.

Rompczyk, K. & Kleinert, C. (2017). *Episode-split biography data in NEPS starting cohort 6: Structure and editing process* (NEPS Survey Paper No. 28). Leibniz Institute for Educational Trajectories, National Educational Panel Study.

Rose, N., von Davier, M., & Nagengast, B. (2017). Modeling Omitted and Not-Reached Items in IRT Models. *Psychometrika, 82*(3), 795–819. https://doi.org/10.1007/s11336-016-9544-7

Shomos, A. (2010). *Links between literacy and numeracy skills and labour market outcomes*. Productivity Commission.

Smith, M. C. (2000). The real-world reading practices of adults. *Journal of Literacy Research, 32*(1), 25–52. https://doi.org/10.1080/10862960009548063

Solga, H., Porsch, P., Ebner, C., & Brzinsky-Fay, C. (2014). The German *vocational education and training system: Its institutional configuration, strengths, and challenge* (WZB Discussion Paper No. SP I 2014–502; WZB Discussion Paper). Wissenschaftszentrum Berlin für Sozialforschung (WZB).

Solheim, O. J., & Lundetræ, K. (2018). Can test construction account for varying gender differences in international reading achievement tests of children, adolescents and young adults? A study based on nordic results in PIRLS, PISA and PIAAC. *Assessment in Education: Principles, Policy & Practice, 25*(1), 107–126. https://doi.org/10.1080/0969594X.2016.1239612

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics, 87*(3), 355–374.

Thums, K., Schwabe, F., Artelt, C., & Wolter, I. (2020). *When do gender differences in reading competence diminish? A longitudinal study of reading competence of women and men from adolescence to young adulthood*. Manuscript submitted for publication.

UNESCO Institute for Statistics (2006). *International standard classification of education: ISCED*. UNESCO Institute for Statistics.

Van Hek, M., Buchmann, C., & Kraaykamp, G. (2019). Educational systems and gender differences in reading: A comparative multilevel analysis. *European Sociological Review*. https://doi.org/10.1093/esr/jcy054

Wicht, A., Müller, N., Haasler, S., & Nonnenmacher, A. (2019). The interplay between education, skills, and job quality. *Social Inclusion, 7*(3), 254. https://doi.org/10.17645/si.v7i3.2052

Wicht, A., Rammstedt, B., & Lechner, C. M. (2020). Predictors of literacy development in adulthood: Insights from a large-scale, two-wave study. *Scientific Studies of Reading*. https://doi.org/10.1080/10888438.2020.1751635

Wicht, A., Durda, T., Krejcik, L., Artelt, C., Grotlüschen, A., Rammstedt, B., & Lechner, C. M. (2020). Low literacy is not set in stone: Longitudinal evidence on the development of low literacy during adulthood. Zeitschrift für Pädagogik. Manuscript accepted for publication.

Widany, S., Christ, J., Gauly, B., Massing, N., & Hoffmann, M. (2019). The quality of data on participation in adult education and training. An analysis of varying participation rates and patterns under consideration of survey design and measurement effects. *Frontiers in Sociology, 4,* 71. https://doi.org/10.3389/fsoc.2019.00071

Windisch, H. (2015). Adults with low literacy and numeracy skills: A literature review on policy intervention, *OECD Education Working Papers* (No. 123), OECD Publishing, https://doi.org/10.1787/5jrxnjdd3r5k-en.

Wu, M., & Nguyen, C. (2019). *The relationship between test length and the number of respondents.* Parkville: University of Melbourne.

Yamamoto, K., Khorramdel, K., & Von Davier, M. (2012a). Chapter 17: Scaling PIAAC cognitive data. In OECD (Ed.), *Technical report of the survey of adult skills (PIAAC)* (2nd edition, p. Chapter 17). OECD.

Yamamoto, K., Khorramdel, K., & von Davier, M. (2012b). Chapter 19: Proficiency scale construction. In OECD (Ed.), *Technical report of the survey of adult skills (PIAAC)* (2nd edition, p. Chapter 19). OECD.

Zabal, A., Martin, S., Klaukien, A., Rammstedt, B., Baumert, J., & Klieme, E. (2013). Grundlegende Kompetenzen der erwachsenen Bevölkerung in Deutschland im internationalen Vergleich. In B. Rammstedt (Ed.), *Grundlegende Kompetenzen Erwachsener im internationalen Vergleich: Ergebnisse von PIAAC 2012* (pp. 31–76). Waxmann.

Zabal, A., Martin, S., Massing, N., Ackermann, D., Helmschrott, S., Barkow, I., & Rammstedt, B. (2014). *PIAAC Germany 2012: Technical report.* Waxmann.

Publisher's Note