Change and Stability in Adults’ Literacy and Numeracy Skills: Evidence from two Longitudinal Large-Scale Surveys

Clemens M. Lechner, Britta Gauly, Alexandra Wicht, & Ai Miyamoto

GESIS – Leibniz Institute for the Social Sciences, Mannheim, Germany

# 

# Abstract

tro

*Keywords*: literacy, numeracy, skills, adult, life course, PIAAC-L?, NEPS?

# Introduction

In today’s knowledge-based and technology-rich societies, literacy (i.e., the ability to understand, use, and interpret written text) and numeracy (i.e., the ability to access, use, and interpret mathematical information) are quintessential skills for the welfare and well-being of both individuals and societies at large. Literacy and numeracy skills are fundamental competencies that are indispensable for handling any type of symbolic verbal or numerical material. As such, they are key prerequisites to acquiring more specific knowledge and skills through such material, not least in increasingly digitized learning environments. It is therefore not surprising that literacy and numeracy are associated with a range of important individual (e.g., income, health, and social participation) and societal outcomes (e.g., economic growth; Hanushek et al., 2015; Hanushek and Woessmann, 2015).

Demographic ageing in most industrialized societies and rapid technological advances imply that an ageing population will be increasingly required to update their skills beyond schooling age and throughout the whole life course, often well into the sixth decade of life. This trend toward lifelong learning poses a number of important questions concerning the development of literacy and numeracy skills in adult age: Can literacy and numeracy skills change during adulthood—or are they already set like plaster after childhood and youth? During which stages of adulthood is skill change most likely to occur, and does this change involve gains or losses? Moreover, how is skill change distributed in the adult population and in major subgroups such as different educational strata? These are not only interesting research questions in their own right, they are also of fundamental importance to policymakers and practitioners interested in promoting lifelong learning. For example, if skills proved to be impervious to change during adulthood, investments in adults’ basic skills such as literacy and numeracy during adulthood might have a low return on investment, and childhood may be a more promising life stage for policies and interventions to focus on (e.g., Cunha & Heckman, 2007). Alternatively, skills may change over time, but gains and losses may be unevenly distributed. In this case, identifying segments of the population that are more likely to experience age-related skill loss than others may aid the development of targeted policies and interventions.

**Previous Evidence on Age Differences in Skills**

Given the policy relevance of these questions about skill development during adulthood, what answers can extant research provide us? Three key insights offered by current evidence are readily summarized (for comprehensive reviews, see Desjardin & Warnke, 2012, and Paccagnella, 2016). First, literacy and numeracy appear to continue to develop across adulthood (lifelong plasticity). Both cross-sectional and the few available longitudinal studies contradict the widespread assumption that skills are set like plaster after childhood (REFERENCES). Second, skill change during adulthood may involve both gains and losses, although typically at different ages (life stage dependency). As is now well documented, the cross-sectional age profile of literacy and numeracy follows an inverted u-shape: On average, literacy and numeracy skills continue to increase throughout the second decade of life, peak at around an age of 30 years, and gradually decline thereafter (Gabrielsen & Lundetræ, 2014; Paccagnella, 2016; Podolsky & Popov, 2013). The resulting age differences in skills are substantial: On average across participating countries, older adults (aged 55–65 years) score around 30 scale points lower on the PIAAC literacy scale (the equivalent of 0.8 *SD*) than young adults aged 25– 34 years (Paccagnella, 2016). Third, change may be unevenly distributed in major socio-demographic subgroups. In particular, those with lower educational attainment often show lower skills compared to those with higher educational attainment (Denny, Harmon, McMahon, & Redmond, 1999; Park & Kyei, 2007; Podolsky & Popov, 2014). Adults who are unemployed (vs. employed) and with low-skilled (vs. high-skilled) occupations are also more likely to have lower skills (Houtkoop & Jones, 1999; Podolsky & Popov, 2014). In addition, we see a general tendency of female adults having relative strength in literacy and of male adults having relative strength in numeracy across the life course (Houtkoop & Jones, 1999; Kline, 2004; Podolsky & Popov, 2014; Satherley & Lawes, 2008). Although adults generally show similar patterns in the skill change across aforementioned socio-demographic subgroups (Reder, 2009; Wilms & Murray, 2007), there is also some research suggesting variability in the magnitude of change across groups. For instance, using PIAAC data, Paccagnella (2016) compared age-related skill change among adults with different levels of educational attainment (i.e., primary, secondary, and territory or above) and found that those with the highest educational attainment experienced the largest skill loss during adulthood.

Although these studies have greatly advanced our knowledge about skill development in adulthood, they cannot conclusively answer the questions posed at the outset. Previous evidence is overwhelmingly based on large-scale cross-sectional surveys such as the Programme for the International Assessment of Adult Competencies (PIAAC) and its predecessors (i.e., the International Adult Literacy Survey in 1994 and 1998 or the Adult Literacy and Life Skills Survey in 2003, 2006, and 2008) ; or on small-scale longitudinal studies based on selective samples such as the longitudinal study of adult learners (LSAL) that focuses on high-school dropouts in the US (Reder, 2009). Cross-sectional studies into the age-related changes in literacy and numeracy skills are limited in that they are unable to disentangle age effects from cohort effects. That is, they are unable to ascertain whether the putative age differences are due to age-related changes or stem from preexisting differences in skills during childhood. Small-scale longitudinal studies based on selective samples, on the other hand, are limited in that their findings may not generalize to the population as a whole. Moreover, by their very nature, these studies cover some subgroups (e.g., high-school dropouts) in a certain life stage (e.g., young adulthood)––but not others subgroups (e.g., the highly educated) and life stages (e.g., old age) that may be of equal interest to policymakers and practitioners. Also, compared to literacy, the life-span development numeracy has received much less attention by prior research, despite arguments in the literature that numeracy skills are gaining in importance on today’s labor markets (REFERENCE).

In order to overcome the limitations of cross-sectional and small-scale longitudinal designs, repeated measures of literacy and numeracy skills are needed. Such data have long been in short supply. Until very recently, there were simply no data sources combining the following desirable features that would allow for complete and robust answers to questions surrounding age-related changes in skills during adulthood: A large and non-selective sample; objective, and high-quality skill assessments; and a repeated measures design.

**The Present Research**

In the present study, we leverage the unique analytical potential of two recent German large-scale assessment surveys that do meet the above criteria: PIAAC-longitudinal (PIAAC-L; Rammstedt et al., 2013), a follow-up to the 2012 Programme for the International Assessment of Adult Competencies (PIAAC) study in Germany; and Starting Cohort 6 from the National Educational Panel Study (NEPS). Both surveys offer repeated, high-quality––and highly comparable––measures of literacy and numeracy spaced three years (PIAAC-L) to six years (NEPS) apart. Combining these data offers us unprecedented opportunities to analyze age-related changes in literacy and numeracy skills during adulthood, allowing us to present what appear to be the most comprehensive descriptive analyses of age-related changes of literacy and numeracy skills during adulthood to date.

With these data, we seek to answer three guiding questions. First, to what extent do literacy and numeracy skills change with age? Second, does the extent of age-related change differ across major socio-demographic groups? Third, to what extent do age differences in skills (as well as potential subgroup variation therein) that we obtain through our repeated-measures designs parallel, or differ from, estimates of age differences obtained in cross-sectional designs (e.g., Paccagnella, 2016)?

We approach these questions from two perspectives on age-related change enabled by the repeated-measures designs.: (1) *relative* change as captured by the correlation between literacy or numeracy skills measured at two time points; and (2) *absolute* change as capture by the change score(for details, see Method). We present descriptive analyses of change both for the total samples as well as in major socio-demographic subgroups defined by age, gender, and educational attainment. These subgroup analyses allow us to detect potential socio-demographic gradients in literacy and numeracy development.

# Method

## Data and Sample

Table 1 shows the basic socio-demographic characteristics of the two samples from PIAAC-L and NEPS at the first measurement occasion.

[Table 1]

**PIAAC/PIAAC-L.** Our first data source is the German PIAAC/PIAAC-L study. PIAAC was conducted in 2012 and provides internationally comparable data on the skills of the working-age population (16–65 years) residing in private households in large number of (mainly OECD) countries (REFERENCE). In Germany, a registry-based sampling design was implemented, in which respondents were randomly sampled from local population registers in randomly selected German municipalities. A total of 5,465 interviews in respondents’ homes were achieved. At the end of the PIAAC interview, all German PIAAC respondents were asked whether they were willing to be re-contacted for a follow-up study (i.e., PIAAC-L) in the future. A total of 3,758 (or 69%) of the original 5,465 PIAAC 2012 respondents consented to being re-interviewed and could be successfully contacted for the first follow-up wave in 2014.

wave

Our analytical sample contains the 2,824 adults for who took the test at both measurement occasions. For the longitudinal analyses we multiplied the total sample weight of PIAAC 2012 with a factor that accounts for non-response in PIAAC-L 2015. s, and weighting or , Zabal et al. (2015) and GESIS, DIW, & LIfBi (2016).

**NEPS.** Our second data source is Starting Cohort 6 (adults) of NEPS (Blossfeld et al., 2011). NEPS is an ongoing large-scale, multi-cohort, longitudinal survey on educational trajectories in Germany (for a detailed description including sampling procedures, see, Blossfeld, Maurice, & Schneider, 2011). Starting Cohort 6 comprised a sample of initially 11,649 adults aged 22–65 years who were interviewed in up to nine annual waves so far.

Literacy and numeracy skills were assessed twice, first in 2010/2011 (henceforth T1) and again in 2016/2017 (T2). Our analytical sample comprises 3,071 respondents who took the literacy test at both occasions and 3,010 respondents for numeracy. We used longitudinal weights where the continuous participation until T2 (wave 9) are computed by means of the longitudinal weight of the previous wave, the probability of being part of the used sample and the likelihood of participating at T2 (for details see XXXX).

## Measures of Literacy and Numeracy

PIAAC/PIAAC-L and NEPS assessed literacy and numeracy in comparable ways, despite some differences in the assessment approach. In both surveys, the tests were low-stakes assessment. Thus, the test scores reflect typical rather than maximal performance. However, the procedures in both surveys were designed to ensure that respondents took the test situation seriously and to guarantee objectivity.

**PIAAC assessment approach.** In PIAAC and PIAAC-L, literacy and numeracy skills were assessed through comprehensive, extensively validated (including cross-nationally) tests. Test tasks were devised by an international commission of eminent scholars. The tasks were designed to reflect tasks relevant to everyday life, which respondents were typically highly motivated to solve. Respondents either took a computer-based assessment (86.2% in our analytical sample in 2012; 88.1% in 2015) or, if they had no experience using a computer, on a paper-pencil version (13.8% in 2012 and 11.8% in 2015). Depending on whether respondents took the computer-based assessment or paper-based assessment, they received different testlets that comprised different subsets of the test tasks. The computer-based assessment was a multistage adaptive testing design that consisted of different literacy and numeracy testlets (each comprising 20 tasks) to which respondents were assigned. The paper-based assessment comprised 20 fixed items for literacy and for numeracy. Interviewers were thoroughly trained for the assessment, they were present while respondents took the tests, and monitored the process. There was no time limit (i.e., tests were not speeded). On average, respondents took about 50 minutes to complete the assessment. For further information on the PIAAC assessment frameworks, see OECD (2012).

The PIAAC/PIAAC-L data distribution provides ten plausible values (PV) ) per respondent and per skill domain and measurement occasion based on item response theory models (IRT). We ran each of our analyses (described below) separately on each of the 10 plausible values and aggregated the results while correcting standard errors (for further details on using plausible values, see Von Davier, Gonzalez, & Mislevy, 2009; Wu, 2005). By using plausible values, covariance-based statistics are corrected for measurement error.

**NEPS assessment approach.** In NEPS, literacy (or, in NEPS terminology, “reading competence”) and numeracy (“mathematical competence”) were assessed through a paper-pencil-based assessment at T1 and a multistage adaptive computer-based assessment at T2. Much akin to PIAAC, the tests were designed to measure skills needed in everyday life. Different to PIAAC, the tests were speeded. The T1 literacy test comprised 32 items, which respondents were asked to complete in 28 minutes. The T1 numeracy test comprised 22 items for which another 28 minutes were available. Respondents were randomly assigned different booklets, and not all respondents took both tests; however, all respondents received all items from a given booklet. At T2, by a multistage design was employed, with respondents receiving one of two booklets of varying difficulty levels depending on their previous performance at T1 (literacy) or their performance in an initial block of tasks at T2 (numeracy). The T2 literacy booklets comprised 27 or 26 items and the numeracy booklet comprised 21 items. For further information on the NEPS assessment framework, see REFERENCE(YEAR).

The NEPS data distribution provides Warm's Mean Weighted Likelihood Estimates (WLE; Warm, 1989) for each skills measure that are linked across measurement occasions so as to be comparable. The WLE are on a logit scale with a mean of zero and unconstrained variance. Note that, in contrast to plausible values in PIAAC-L, the WLE are manifest scores. They contain measurement error that will attenuate correlations (i.e., underestimate the true correlations), although means will be unbiased.

**Comparability of PIAAC and NEPS.** In wave 2 of PIAAC-L (2015), a linking study was conducted to compare the PIAAC and NEPS tests of literacy and numeracy (Carstensen et al., in preparation). For this purpose, participants received different test versions, some of which contained items from both PIAAC and NEPS. This allows to jointly scale the test results using item response theory (IRT) methods. Results from this linking study show that the PIAAC and NEPS test correlate highly: The PIAAC literacy test correlates at *r* = . 87 with the NEPS literacy test, and the PIAAC numeracy test correlates at *r* = .90 with the NEPS numeracy test. Moreover, the correlations between the literacy and numeracy tests within each study were also very similar: PIAAC literacy correlated with PIAAC numeracy at *r* = .87, and NEPS literacy correlated with NEPS numeracy at *r* = .87. This pattern of correlations suggests that, despite the differences in the assessment approaches of both surveys, the tests yield comparable results and measure largely the same latent constructs.

## Descriptive Indices of Change and Stability

As noted at the outset, we compute several complementary indices to gauge the extent of relative and absolute change in literacy and numeracy skills, described below. We compute each of these three indices for the total sample as well as separately by age, educational attainment, and gender. This allows us to detect potential socio-demographic gradients in skill development.

**Relative change.** Relative (or rank-order) changerefers to change in individuals’ relative standing in the skill distribution across two measurement occasions. Its measure is the correlation between the tests at two time points (*r*T1, T2). In PIAAC/PIAAC-L, this correlation is based on 10 plausible values (PV) per person and implicitly corrected for measurement error. In NEPS, the correlation is based on the WLE scores per person, which still contain measurement error; thus, NEPS correlations provide a lower bound estimate of the true correlation. Apart from these differences, , *r*T1, T2 is a standardized effect size measure that can be readily compared within and across studies.

**Absolute change.** The most straightforward measure of differences in skills between two measurement occasions is *intra-individual change*, captured by the change score (ΔT2, T1).[[1]](#footnote-1) Whereas *r*T1, T2 refers to how stable individuals’ relative standing in the skill distribution is across two measurement occasions, ΔT2, T1 refers to the how much individuals’ skills change over time in absolute terms. When aggregating ΔT2, T1 across all individuals in the sample or a subgroup, we obtain an estimate of *mean-level change*.

Absolute change is largely independent of relative change: Even if there is hardly any rank-order change (e.g., *r*T1, T2 > .90), this does not preclude the possibility of mean-level change—as long as skill gains or losses are very similar in size across individuals (i.e., ΔT2, T1  has little variability). In contrast to rank-order change (*r*T1, T2), ΔT2, T1 is in the *raw* metric of the T1 and T2 test scores (0–500 points in PIAAC; logits in NEPS). Therefore, to enhance comparability within and across studies, we also report Cohen’s *dz*, which expresses the mean difference ΔT2, T1 in units of the population standard deviationof the test. As is conventional, we used the pooled standard deviation from both time points as an estimate of :

|  |  |
| --- | --- |
|  |  |

Additionally, to gauge how many individuals experienced statistically significant intra-individual changefrom one measurement occasion to the next, we computed the reliable change index (RCI; Jacobson & Truax, 1991; Christensen & Mendoza, 1986) and its closest equivalent for item response theory (IRT) models, the *Z*-test (see Brouwer et al., Meijer, Zevalking, 2013). RCI and *Z* express change between two measurement occasions relative to the standard error of measurement of the difference score. The RCI is calculated as

|  |  |
| --- | --- |
|  | (2) |

where refers to the competence measure in and refers to the competence measure in . is the standard deviation of the competence measure in and is the correlation between the two competence measures (also see, Brouwer, 2013).

As described above, PIAAC not only contains one competence score for each person, but 10 plausible values. For the calculation of the RCI while taking into account the plausible values, we computed Equation (2) ten times with one of the values each and averaged the ten resulting RCIs.

NEPS formula

RCI and *Z* asymptotically follows a standard normal distribution. Thus, values greater than ±1.96 are typically interpreted as statistically significant change, as it is unlikely to occur by chance in more than 5% of cases.

**Subgroup analyses.** Our subgroup analyses required that we split the age and education variables in a way that was (a) theoretically meaningful; (b) fine-grained enough to detect potential differences while (c) still resulting in subgroups large enough to ensure stable estimates in both NEPS and PIAAC-L. Based on these criteria, we chose to code age in four groups: young adulthood (18–29 years), prime age (30–44 years), mid adulthood (45–35 years) and older adulthood (55 years and older). We chose to code educational attainment in three groups according to the International Standard Classification of Education (ISCED 1997): lower or upper secondary (“low”; ISCED levels 0–3), post-secondary or tertiary vocational education (“intermediate”; ISCED levels 4, and 5B), and tertiary academic education (“high”; ISCED levels 5A and 6).

# Results

## Results from PIAAC/PIAAC-L: Three-Year Change in Literacy and Numeracy

**Relative change.** Table 2 (left column) shows the correlations of literacy and numeracy over three years in PIAAC-L, which serve as our measure of relative (i.e., rank-order) change. The upper part of the table shows the estimates for the total population, whereas the lower part shows estimates for each of the socio-demographic subgroups.

The very high correlation (*r*T1, T2) of literacy over three years in the total population indicates that individuals’ relative standing in the skill distribution remains highly stable over time. Still, the correlation is still far from unity, and some rank-order changes do occur. The three-year stability for numeracy is similar to that of literacy. The correlations imply that the skill measures at the two measurement occasions spaced three years apart share 72% (literacy) and 67% (numeracy) of their variance.

[Table 2]

When estimating the three-year rank-order change separately by major socio-demographic subgroups (Table 2), Quite a few differences emerge, most of which are statistically significant at a reasonably strict level of α <. 001: Both literacy and numeracy are significantly more stable in young adulthood (18–29 years) than in old adulthood (55+ years), with the other age groups falling in between. Literacy and numeracy are also significantly more rank-order stable among the lower-educated than among the higher-educated and among. Both skills are significantly more stable among the male adults than among the female adults.

For a more intuitive interpretation of the stability estimates for the total population, Figure 1 displays these three-year correlations for literacy (Panel A) and numeracy (Panel B) graphically. Each dot in the cloud represents one individual’s *z*-standardized skills measured at T1 (X axis) and T2 (Y axis). The gray line is the actual correlation. The red line represent a (hypothetical) perfect correlation (*r*T1, T2 = 1.00). A larger angle between the gray and red line indicate a higher the rank-order stability. For literacy and numeracy, the angle is small, indicating that the actual correlation is close to a perfect one. The dot cloud is very densely populated around the gray line, with only few outliers. For numeracy, there are a few more individuals whose T2 scores deviate more strongly from their T1 scores, yet the overall pattern is similar to that of literacy and indicative of high rank-order stability.

[Figure 1]|

**Absolute change.** Next, we examine intra-individual and mean-level change (ΔT2, T1) in literacy and numeracy. Table 3 (left column) shows means and standard deviations of change (ΔT2, T1) in literacy and numeracy over three years for the total population and for each socio-demographic subgroup, along with the respective Cohen’s *d* as an effect size measure.

[Table 3]

As can be seen, there are no mean-level changes in literacy in the total adult population over three years. However, as the standard deviations of the change scores suggest, there is considerable variation in change around the mean. To highlight this variation, Figure 2 shows the distribution of the change scores (ΔT2, T1) over a three-year period for literacy (left panel) and numeracy (right panel). As can be seen, this distribution is normal, indicating that gains and losses are equally likely. That is, the absence of change in the average change score on average actually conceals considerable heterogeneity in the extent of change. A substantial share of respondents, namely XX%, experience skill gains or losses that exceeds one standard deviation (indicated by the dotted line).

[Figure 2]

Table 3 also reveals some differences in mean-level change across socio-demographic subgroups. Specifically, young adults (18–29 years) experience non-negligible average gains in both literacy and numeracy. Moreover, the most highly respondents experience average gains in literacy (but not numeracy) over three years.

What share of individuals experience change in literacy and numeracy that can be considered statistically significant? According to the RCI values shown in Table 4 (right column), XX% respondents experience gains in literacy that exceeds

## Results from NEPS SC6: Six-Year Change in Literacy and Numeracy

**Relative change.** Turning to relative change in literacy and numeracy over six years in NEPS (Table 2, right column), in the total sample, we find relatively high correlation (*r*T1, T2) of both literacy and numeracy. As expected, six-year stabilities for literacy and numeracy are considerably lower than the three-year consistencies in PIAAC-L, and lower than what one would expect by simply extrapolating the three-year correlation from PIAAC-L (.85 ✕ .85 = .72 for literacy; .82 ✕ .82 = .67 for numeracy). The correlations imply that the two measures of skills spaced six years apart share only 38% of their variance for literacy and 50 % of variance for numeracy, which as expected, is much lower than that for skills over three years.

Akin to PIAAC-L, we also found some differences in the rank-order stability between different socio-demographic subgroups in NEPS SC6. The rank-order stability of literacy and numeracy is significantly higher in young adulthood (18–29 years) and lower in old adulthood (55+ years). The rank-order stability of numeracy (but not of literacy) is also significantly higher among the higher-educated than the lower-educated. There was no significant difference in the rank-order stability for males and females.

Figure 3 shows the graphical representations of six-year correlations for literacy and numeracy for the total adult population. Notice that, compared to the three-year stabilities in Figure 1, the angle between the (hypothetical) perfect correlation in red line and the actual correlation in gray line is now larger for both literacy and numeracy, indicating greater deviations from perfect stabilities. Moreover, the dot cloud is much more dispersed for both literacy and numeracy compared to PIAAC-L. Also notice a small number of individuals who concentrate at the upper end of the literacy and numeracy distribution (i.e., in the upper right corner of the figure). This pattern suggests the presence of ceiling effects in the NEPS literacy and numeracy tests for these individuals.

[Figure 3]

## Absolute change. With regard to the intra-individual and mean-level change (ΔT2, T1) in literacy and numeracy (Table 3, right column), we found no substantial mean-level changes in literacy and numeracy in the total sample over six years. There is nevertheless considerable variation in change around the mean for both skills. The distribution of the change scores (ΔT2, T1) over a six-year period for literacy (above) and numeracy (below) can be found in Figure 4. The distribution is normal showing equal likelihood of gains and losses in both skills. There are even a higher percentage of individuals clustering in the middle (around zero) for numeracy compared to literacy, indicating that more individuals experienced little change in numeracy than in literacy.

[Figure 4]

Similar to PIAAC-L, we see some differences in mean-level change across socio-demographic subgroups (Table 3). Again, young adults (18–29 years) experience average gains in literacy and to a lesser extent, numeracy. Moreover, most low-educated individuals experience average gains in literacy (but not numeracy) over six years. In comparison, most high-educated individuals experience average losses in both literacy and numeracy over six years.

## Table 4 (right column): RCIs

# Conclusion

# lower stability for longer time period (NEPS) indicates that there is room for skill development in adulthood, it just needs time

* even in higher age groups, there are some individuals who experience gains in skills (results from RCI, Table 5)
* Matthew effect not supported, as lowest educational groups experience skill gains (especially in NEPS)
* limitations
  + longer panel data necessary
  + broader set of skills relevant for today’s technology driven labor market
* further research (from Bratislava Keynote)
  + Whence the **variability in skill development**?
  + General slowing hypothesis / biological ageing?
  + Practice engagement effects?
  + What are the **consequences of skill loss**?
  + For life outcomes (income, social participation, well-being, etc.)?
  + For learning ability?
  + Linear functions vs. critical thresholds?
  + What are the **returns on investments** in adults’ literacy and numeracy skills?
  + How and when to deliver interventions?
  + For what subgroups?

# References

Denny, K., Harmon, C., McMahon, D., & Redmond, S. (1999). Literacy and education in Ireland. *The Economic and Social Review*, *30*(3), 215-226. doi:10.1920/wp.ifs.2000.0009

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, 72*. Paris: OECD Publishing. <http://dx.doi.org/10.1787/5k9csvw87chk-en>

GESIS, DIW, LIfBi (2016). PIAAC-L 2014 and 2015: Preliminary Notes to the User.

Hanushek, E. A., & Woessmann, L. (2015). The knowledge capital of nations: Education and the economics of growth.

Hanushek, E. A., Schwerdt, G., Wiederhold, S., & Woessmann, L. (2015). Returns to skills around the world: Evidence from PIAAC. European Economic Review, 73, 103-130.

Houtkoop, W., & Jones, S. (1999). Adult numeracy: an international comparison. In *Mathematics as Part of Lifelong Learning: Proceedings of the Fifth International Conference of Adults Learning Maths–A Research Forum, ALM-5, London: Goldsmiths College, University of London in association with ALM* (pp. 32-40).

Kline, T. J. (2004). Gender and language differences on the test of workplace essential skills: Using overall mean scores and item-level differential item functioning analyses. *Educational and Psychological Measurement*, *64*(3), 549-559. https://doi.org/10.1177/0013164403258444

Paccagnella, M. (2016). Age, ageing and skills: Results from the survey of adult skills. OECD Education Working papers No. 132, OECD Publishing, Paris. http://dx.doi.org/10.1787/5jm0q1n38lvc-en

Park, H., & Kyei, P. (2007). Literacy gaps by educational attainment: A comparative study of 19 countries. PARC Working Paper Series,WPS 07-12.

Podolskiy, O., & Popov, D. (2014). [The first assessment of adult competencies in Russia](https://ideas.repec.org/a/nos/voprob/2014i2p82-108.html). [*Educational Studies*](https://ideas.repec.org/s/nos/voprob.html)*, Higher School of Economics* (2),82-108.

Reder, S. (2009). The development of adult literacy and numeracy in adult life. In S. Reder and J. Bynner, J. (Eds), Tracking Adult Literacy and Numeracy Skills: Findings from Longitudinal Research (pp. 59-84). Routledge: New York.

Satherley, P., & Lawes, E. (2008a). The Adult Literacy and Life Skills (ALL) Survey: Gender, ethnicity and literacy. Wellington: Ministry of Education.

Scharl, A., Carstensen, C. H., & Gnambs, T. (2020).Estimating Plausible Values with NEPSData: An Example Using Reading Competence in Starting Cohort 6(NEPS Survey Paper No.∞). Bamberg: Leibniz Institute for Educational Trajectories, National Educational Panel Study. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/Kompetenzen/NEPSscaling\_preprint.pdf

Warm, T. A. (1989). Weighted likelihood estimation of ability in Item Response Theory. Psychometrika, 54, 427-450.

Willms, D., & Murray. T. S. (2007). Gaining and losing literacy skills over the life course, Statistics Canada, Ottawa.

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

1. Mean-level stability of skills can be estimated from cross-sectional designs, such as by comparing skills between different age groups (e.g,. Paccagnella, 2016). However, as noted earlier, cohort and period effects may confound putative age effects in such between-subjects designs. [↑](#footnote-ref-1)