

FDZ-Datenreport

Documentation of labour market data

05/2012

EN

Working and Learning in a Changing World

Part VII: Description of the ALWA literacy and numeracy
data (ALWA-LiNu)

Oliver Wölfel
Corinna Kleinert



Working and Learning in a Changing World

Part VII: Description of the ALWA literacy and numeracy data (ALWA-LiNu)

Oliver Wölfel (Institute for Employment Research)

Corinna Kleinert (Institute for Employment Research)

Die FDZ-Datenreporte beschreiben die Daten des FDZ im Detail. Diese Reihe hat somit eine doppelte Funktion: zum einen stellen Nutzerinnen und Nutzer fest, ob die angebotenen Daten für das Forschungsvorhaben geeignet sind, zum anderen dienen sie zur Vorbereitung der Auswertungen.

FDZ-Datenreporte (FDZ data reports) describe FDZ data in detail. As a result, this series has two aims: first, users can ascertain whether the data are suitable for their research; second, the reports can be used to prepare analyses.

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 5 |
| 2 | Data description | 6 |
| 2.1 | <i>Outline</i> | 6 |
| 2.2 | <i>Test implementation</i> | 7 |
| 3 | Description of variables | 8 |
| 3.1 | <i>Technical information</i> | 8 |
| 3.1.1 | Personal ID (caseid) | 8 |
| 3.1.2 | Test administration status (teststat) | 8 |
| 3.1.3 | Booklet version (bookid) | 8 |
| 3.2 | <i>Weights</i> | 9 |
| 3.2.1 | Benchmark weight (bmweight) | 9 |
| 3.2.2 | Jackknife weights (replicate weights) (jack1-jack30) | 9 |
| 3.3 | <i>Proficiency estimates</i> | 10 |
| 3.3.1 | Maximum Likelihood estimates (ML) for prose literacy (ml_l_est) | 10 |
| 3.3.2 | Maximum Likelihood standard error (ML) for prose literacy (ml_l_se) | 10 |
| 3.3.3 | Maximum Likelihood estimates (ML) for numeracy (ml_n_est) | 10 |
| 3.3.4 | Maximum Likelihood standard error (ML) for numeracy (ml_n_se) | 10 |
| 3.3.5 | Expected A Posteriori estimates (EAP) for prose literacy (eap_l_est) | 11 |
| 3.3.6 | Expected A Posteriori standard deviation (EAP) for prose literacy (eap_l_sd) | 11 |
| 3.3.7 | Expected A Posteriori estimates (EAP) for numeracy (eap_n_est) | 11 |
| 3.3.8 | Expected A Posteriori standard deviation (EAP) for numeracy (eap_n_sd) | 11 |
| 3.3.9 | Plausible Values for literacy (pv_l1-pv_l5) | 11 |
| 3.3.10 | Plausible Values for numeracy (pv_n1-pv_n5) | 11 |
| 3.4 | <i>Original test items</i> | 12 |
| 3.4.1 | Core booklet (core1-core10) | 12 |
| 3.4.2 | Main booklet items block 1 literacy (b1q1-b1q12) | 12 |
| 3.4.3 | Main booklet items block 2 literacy (b2q1-b2q13) | 12 |
| 3.4.4 | Main booklet items block 3 numeracy (b3q1-b3q11) | 12 |
| 3.4.5 | Main booklet items block 4 numeracy (b4q1-b4q11) | 12 |
| 4 | Data access | 14 |
| | References | 15 |

Abstract

This report describes the variables of ALWA-LiNu data. The IAB study ALWA ('Working and Learning in a Changing World') was conducted in 2007/08 and contains detailed retrospective life course data of 10,404 individuals living in Germany born in between 1956 and 1988. ALWA-LiNu provides information of 3,980 study participants on their cognitive skills in two domains, prose literacy and numeracy.

Zusammenfassung

Der Datenreport beschreibt die Variablen des ALWA-LiNu-Datensatzes. Die IAB-Studie ALWA ('Arbeiten und Lernen im Wandel') wurde 2007/08 erhoben und enthält detaillierte retrospektive Lebensverlaufsdaten von 10.404 in Deutschland lebenden Personen der Geburtskohorten 1956 bis 1988. ALWA-LiNu liefert Informationen von 3.980 Studienteilnehmern zu ihren kognitiven Kompetenzen in zwei Domänen, dem Verständnis von Fließtexten (prose literacy) und Zahlen (numeracy).

We would like to thank Alexandra Schmucker (FDZ) for her support writing this report.

Keywords: ALWA, competence tests, cognitive competencies, literacy, numeracy

1 Introduction

The ALWA-LiNu dataset contains a subsample of cases from the ALWA survey. The IAB study ALWA ('Working and Learning in a Changing World') was conducted in 2007/08 and contains detailed retrospective life course information of 10,404 individuals living in Germany born in between 1956 and 1988 (for further information, see Antoni et al. 2010). The ALWA life course questionnaire was conducted in computer-assisted telephone interview (CATI). At the end of the telephone interview, participants were asked for consent to participate in a personal interview taking place some days later.

With those persons who gave their consent and agreed in participating, paper-and-pencil personal interview (PAPI) were conducted. These interviews mainly contained cognitive competence tests in two domains, prose literacy and numeracy (for detailed information on study design, instruments and implementation, see Kleinert et al. 2012). Valid test data was gathered for a subsample of 3,980 ALWA study participants.

Test data of this subsample is provided for users in a new dataset named ALWA-LiNu. In this data report the variables of the ALWA-LiNu dataset are described. It contains four blocks of information:

- technical information on booklets and test completion,
- weights,
- proficiency estimators and
- original test items.

A detailed description and discussion of weights and proficiency estimators as well as practical advice how to use them is provided in Kleinert et al. (2012).

2 Data description

2.1 Outline

Content characteristics:

| | |
|------------------------|---|
| Topics | Item-test scores |
| Data unit | Resident population of the FRG, birth cohorts 1956-1988 |
| Number of cases | 3,980 individuals |
| Period covered | August 2007- August 2008 |
| Data collection | August 2007-August 2008 |

Methodological characteristics:

| | |
|-------------------------------------|---|
| Method of data collection | Face-to-face paper-and-pencil interview (PAPI) |
| Institutions involved | Institute for Employment Research (IAB), Nuremberg Infas – Institute for Applied Social Sciences, Bonn ETS – Educational Testing Service, Princeton, USA Statistics Canada, Ottawa, Canada |
| Frequency of data collection | Once |
| Time reference | Date of interview / test |
| File format | Stata, SPSS |
| File size | 1 MB (Stata), 2 MB (SPSS) |
| File architecture | ALWA-LiNu.dta (Stata) ALWA-LiNu.sav (SPSS) |

Data access:

| | |
|--------------------------------|---------------------|
| Data access | Scientific use file |
| Degree of anonymisation | Factually anonymous |
| Sensitive variables | None |

The dataset described in this document is available for use by professional researchers. Further information can be found on the website <http://fdz.iab.de>.

2.2 Test implementation

Two task booklets, the core and the main booklet, were employed in the face-to-face interview. In order to maintain reliability and validity of the items in further surveys to derive literacy and numeracy proficiencies the wording of the different items cannot be provided. Examples for similar items can be found in StatCan & OECD 2005.

The short **core booklet** contained 5 literacy and 5 numeracy items (variables *core1-10*). These tended to be the easiest items in the whole item pool of the ALWA test. The aim of the core booklet was to identify respondents with very low literacy and numeracy competencies in order to avoid forcing them to answer a long test with complicated tasks. In order to do so, the interviewers had to score the respondents' core booklet answers by marking and summing up all correct answers. At maximum respondents could achieve 10 points, at minimum 0 points. If the sum was 4 points (the cut-off point) or below the interview was finished. If the sum was higher, the interview was continued with the main booklet. Information on the respondents' result regarding this aim is found in the ALWA-LiNu data in the variable *teststat*.

The **main booklet** was implemented in four versions signed by four different colors. For compiling the versions, the prose literacy and the numeracy items were divided by two for each scale, resulting in four blocks of cognitive items: literacy block 1 (variables *b1q1-12*), literacy block 2 (variables *b2q1-13*), numeracy block 3 (variables *b3q1-11*) and numeracy block 4 (variables *b4q1-11*).

The design of the four main booklet versions is presented in Table 1. Each of the blocks appears once in every position. Information on the booklet version of every respondent is provided in the variable *bookid*.

Table 1. ALWA Main Booklet Design

| | Booklet A | Booklet B | Booklet C | Booklet D |
|---------------------|------------------|------------------|------------------|------------------|
| First block | Literacy block 1 | Numeracy block 3 | Literacy block 2 | Numeracy block 4 |
| Second block | Numeracy block 3 | Literacy block 1 | Numeracy block 4 | Literacy block 2 |
| Third block | Literacy block 2 | Numeracy block 4 | Literacy block 1 | Numeracy block 3 |
| Fourth block | Numeracy block 4 | Literacy block 2 | Numeracy block 3 | Literacy block 1 |

3 Description of variables

3.1 Technical information

3.1.1 Personal ID (caseid)

| | |
|----------------|--|
| Variable label | Person number |
| Variable name | caseid |
| Value label | |
| Description | Unique identification number of respondent |

3.1.2 Test administration status (teststat)

| | |
|----------------|---|
| Variable label | Test administration status |
| Variable name | teststat |
| Value label | 1 Completely realized 2 Aborted due to excess length 3 Core booklet not passed |
| Description | Numerical variable which indicates whether the test was completely realized, aborted because it took too long or if the main test was not administered because the respondent performed to low in the core booklet. |

3.1.3 Booklet version (bookid)

| | |
|----------------|--|
| Variable label | Booklet version number |
| Variable name | bookid |
| Value label | 1 Booklet A 2 Booklet B 3 Booklet C 4 Booklet D |
| Description | Numerical variable which indicates the booklet version of every respondent |

3.2 Weights

3.2.1 Benchmark weight (bmweight)

| | |
|----------------|---|
| Variable label | Benchmark weight |
| Variable name | bmweight |
| Value label | |
| Description | Survey weight which adjusts the marginal distribution of the sample to the distribution known from official statistical data (Statistical Office). The weight provided here adjusts for selection bias, but is does not extrapolate for the total population. |

3.2.2 Jackknife weights (replicate weights) (jack1-jack30)

| | |
|----------------|--|
| Variable label | Jackknife weight 1-30 |
| Variable name | jack1-jack30 |
| Value label | |
| Description | 30 variables containing the Jackknife weights on the basis of 30 replicates, needed in order to be able to estimate standard errors without knowing the Primary Sample Units (PSUs). |

3.3 Proficiency estimates

3.3.1 Maximum Likelihood estimates (ML) for prose literacy (ml_l_est)

| | |
|----------------|--|
| Variable label | ML literacy: estimate |
| Variable name | ml_l_est |
| Value label | --- |
| Description | Contains a point estimate for prose literacy using the Maximum Likelihood method |

3.3.2 Maximum Likelihood standard error (ML) for prose literacy (ml_l_se)

| | |
|----------------|---|
| Variable label | ML literacy: standard error |
| Variable name | ml_l_se |
| Value label | --- |
| Description | Contains the standard error of the ML prose literacy estimate |

3.3.3 Maximum Likelihood estimates (ML) for numeracy (ml_n_est)

| | |
|----------------|--|
| Variable label | ML numeracy: estimate |
| Variable name | ml_n_est |
| Value label | |
| Description | Contains a point estimate for numeracy using the Maximum Likelihood method |

3.3.4 Maximum Likelihood standard error (ML) for numeracy (ml_n_se)

| | |
|----------------|---|
| Variable label | ML numeracy: standard error |
| Variable name | ml_n_se |
| Value label | |
| Description | Contains the standard error of the ML numeracy estimate |

3.3.5 Expected A Posteriori estimates (EAP) for prose literacy (eap_l_est)

| | |
|----------------|---|
| Variable label | EAP literacy: estimate |
| Variable name | eap_l_est |
| Value label | |
| Description | Contains the mean of the posterior distribution of prose literacy |

3.3.6 Expected A Posteriori standard deviation (EAP) for prose literacy (eap_l_sd)

| | |
|----------------|--|
| Variable label | EAP literacy: standard deviation |
| Variable name | eap_l_sd |
| Value label | |
| Description | Contains the standard deviation of the prose literacy posterior distribution |

3.3.7 Expected A Posteriori estimates (EAP) for numeracy (eap_n_est)

| | |
|----------------|---|
| Variable label | EAP numeracy: estimate |
| Variable name | eap_n_est |
| Value label | |
| Description | Contains the mean of the posterior distribution of numeracy |

3.3.8 Expected A Posteriori standard deviation (EAP) for numeracy (eap_n_sd)

| | |
|----------------|--|
| Variable label | EAP numeracy: standard deviation |
| Variable name | eap_n_sd |
| Value label | |
| Description | Contains the standard deviation of the numeracy posterior distribution |

3.3.9 Plausible Values for literacy (pv_l1-pv_l5)

| | |
|----------------|---|
| Variable label | Literacy: plausible value 1-5 |
| Variable name | pv_l1 pv_l2 pv_l3 pv_l4 pv_l5 |
| Value label | |
| Description | Five variables containing Plausible Values estimates for prose literacy |

3.3.10 Plausible Values for numeracy (pv_n1-pv_n5)

| | |
|----------------|---|
| Variable label | Numeracy: plausible value 1-5 |
| Variable name | pv_n1 pv_n2 pv_n3 pv_n4 pv_n5 |
| Value label | |
| Description | Five variables containing Plausible Values estimates for numeracy |

3.4 Original test items

3.4.1 Core booklet (core1-core10)

| | |
|----------------|---|
| Variable label | Core: question 1-10 [literacy/numeracy] |
| Variable name | core1- core10 |
| Value label | 0 Incorrect answer 1 Correct answer 3 No answer |
| Description | 10 numerical variables containing the answers to the tasks in the core booklet, mixed prose literacy and numeracy tasks |

3.4.2 Main booklet items block 1 literacy (b1q1-b1q12)

| | |
|----------------|--|
| Variable label | Block 1 literacy: question 1-12 |
| Variable name | b1q1-b1q12 |
| Value label | 0 Incorrect answer 1 Correct answer 3 No answer |
| Description | 12 numerical variables containing the answers to the tasks in literacy block 1 of the main booklet |

3.4.3 Main booklet items block 2 literacy (b2q1-b2q13)

| | |
|----------------|---|
| Variable label | Block 2 literacy: question 1-13 |
| Variable name | b2q1-b2q13 |
| Value label | 0 Incorrect answer 1 Correct answer 3 No answer |
| Description | 13 numerical variables containing the answers to the tasks literacy block 2 of the main booklet |

3.4.4 Main booklet items block 3 numeracy (b3q1-b3q11)

| | |
|----------------|--|
| Variable label | Block 3 numeracy: question 1-11 |
| Variable name | b3q1-b3q11 |
| Value label | 0 Incorrect answer 1 Correct answer 3 No answer |
| Description | 11 numerical variables containing the answers to the tasks in the numeracy block 3 of the main booklet |

3.4.5 Main booklet items block 4 numeracy (b4q1-b4q11)

| | |
|----------------|---------------------------------|
| Variable label | Block 4 numeracy: question 1-11 |
|----------------|---------------------------------|

| | |
|---------------|--|
| Variable name | b4q1-b4q11 |
| Value label | 0 Incorrect answer 1 Correct answer 3 No answer |
| Description | 11 numerical variables containing the answers to the tasks in the numeracy block 4 of the main booklet |

4 Data access

The dataset ALWA-LiNu is available as extension file for the scientific use file of the ALWA survey data as well as the ALWA survey data linked to administrative data of the IAB (ALWA-ADIAB) via the Research Data Centre (FDZ) of the Federal Employment Agency at the Institute for Employment Research. It contains only the variables described in the previous section for 3,980 respondents with valid test data. In order to gain background information on the test participants ALWA-LiNu has to be merged with the ALWA survey data using the key variable 'caseid' (for the CATI codebook, see Matthes & Trahms 2010).

References

Antoni, Manfred; Drasch, Katrin; Kleinert, Corinna; Matthes, Britta; Ruland, Michael; Trahms, Annette (2010): Working and learning in a changing world * Part I: Overview of the study - March 2011 (Second, updated version). (FDZ Methodenreport, 05/2010 (en)), Nürnberg

Kleinert, Corinna; Yamamoto, Kentaro; Wölfel, Oliver; Gilberg, Rainer (2012): Literacy and Numeracy Skills – Test Design, Implementation, Scaling and Statistical Models for Proficiency Estimation. FDZ-Methodenbericht Nr. 10/2012, Nürnberg.

Matthes, Britta; Trahms, Annette (2010): Working and Learning in a Changing World Part II - Codebook. (FDZ Datenreport, 02/2010 (de)), Nürnberg

Statistics Canada [StatCan]; Organisation for Economic Co-operation and Development [OECD] (Hrsg.) (2005): Learning a Living. First Results of the Adult Literacy and Life Skills Survey. StatCan/OECD: Ottawa and Paris.

Imprint

FDZ-Datenreport 05/2012

Publisher

The Research Data Centre (FDZ)
of the Federal Employment Agency
in the Institute for Employment Research
Regensburger Str. 104
D-90478 Nuremberg

Editorial staff

Stefan Bender, Dagmar Theune

Technical production

Dagmar Theune

All rights reserved

Reproduction and distribution in any form, also in parts,
requires the permission of FDZ

Download

http://doku.iab.de/fdz/reporte/2012/DR_05-12_EN.pdf

Internet

<http://fdz.iab.de/>

Corresponding author:

Oliver Wölfel
Institute for Employment Research,
Regensburger Str. 104,
D-90478 Nuremberg
Email: Oliver.Woelfel@iab.de