March 2021

# German Ageing Survey (DEAS):
## User Manual SUF DEAS 2011, Version 2.2

*Heribert Engstler, Nicole Hameister,*
*Ruta Daktariunaite & Beate Schwichtenberg-Hilmert*

# CONTENTS

## PRELIMINARY REMARKS: CHANGES IN SUF VERSION 2.2

Apart from minor label editions, SUF 2.2 of DEAS 2011 contains the following changes compared to the previous version:

- Sum score values of variable *depressiv_11* are generated differently. It's not necessary any longer that all 15 items have valid values to generate a sum score. Instead, a mean score is generated when at least 8 items have valid values. Mean score is then multiplied by 15 to get the sum score. This avoids to exclude respondents with non-response on 1 to 7 items.

## PRELIMINARY REMARKS: CHANGES IN SUF VERSION 2.1

Compared to the previous version, version 2.1 of the Scientific Use File of the DEAS survey 2011 contains the following changes:

- The variables to the relative income positon, to the income poverty and the income wealth have been upgraded, based on new reference-numbers (see chapter 8.15)

- New construct variables were made related to nationality, migration background and education contain data for all samples (baseline and panel respondents). For panel respondents the characteristics have been adopted from the first measuring time

- The longitudinal weighting has been updated. From now on the Scientific Use File (SUF) contains an integrated cross-sectional weighting for all interviewees up to the age of 90 (see chapter 2).

## 1   THE GERMAN AGEING SURVEY (DEAS) – A STUDY ON THE SECOND HALF OF LIFE

The German Ageing Survey (DEAS) is a nationwide representative cross-sectional and longitudinal survey of the German population aged over 40. It is funded by the Federal Ministry for Family Affairs, Senior Citizens, Women and Youth (BMFSFJ). The German Centre of Gerontology in Berlin (DZA) is responsible for the conduct and ongoing development of the study. Sampling and fieldwork for all waves (1996-2020) has been carried out by the Bonn-based Institute for Applied Social Sciences (infas). The primary goal of the DEAS survey program is to provide a representative national database containing information describing the living conditions of the country's middle-aged and older population and to study diversity within the older section of the population, the process of ageing as it affects individuals and processes of social change as they relate to old age and ageing.

The DEAS covers a wide range of topics. The data obtained provide information on socioeconomic and demographic attributes as well as household composition, housing, family structure, social networks, psychological resources, attitudes as well as and physical

and mental health. The comprehensive examination of people in mid- and older adulthood provides micro data for use both in social and behavioral scientific research and in reporting on social developments. The data thus provides a source of information for decision-makers, the general public and for scientific research.

The DEAS applies a cohort-sequential design, which allows the users to analyze societal trends and individual trajectories (embedded inside societal trends) and to disentangle age effects from cohort effects. The first DEAS survey wave took place in 1996, further waves followed in 2002, 2008, 2011, 2014, 2017 and 2020. The 4th wave of the survey in 2011 considered a panel sample of study participants, who had entered the DEAS earlier.

Microdata of the German Ageing Survey (DEAS) are available free of charge to scientific researchers for non-profitable purposes. The Research Data Centre (FDZ-DZA) provides access and support to scholars interested in using DEAS data for their research. Data and documentations from completed DEAS waves are available by the FDZ-DZA. However, for reasons of data protection, signing a data distribution contract is required before data can be obtained.


## 1.1   Design and sampling

The DEAS uses a combination of cross-sectional and longitudinal samples. Since 1996 every six years a new baseline sample of community dwelling 40- to 85 year-olds is drawn up. The baseline samples are stratified by age group, sex, and place of residence (East- or West-Germany) and drawn from the registration office. Using cross-sectional weights to correct for this the DEAS baseline samples are nationally representative for adults aged from 40 to 85 years. Participants are interviewed personally by interviewers. Since wave 2 (2002), all panel-willing participants from the baseline samples are tracked. Participants of the foreigner sample in 2002 were excluded from the panel. After the 3rd wave in the year 2008, panel members are interviewed again after three years already. Therefore, in 2011 there was only a panel sample.

The target population in 1996 was defined as German citizens residing in the community. In 2002, a comparable sample of German citizens was drawn up; in addition, a random sample of non-German citizens residing in Germany was set up separately. Since 2008, German and non-German citizens have been drawn up together from the population residing in the community in Germany.

Baseline participants, who gave written consent were re-contacted for further waves of data collection. Panel attrition is high in the first re-interview but attenuates in subsequent follow-ups. Information of a single DEAS wave is put together in a Scientific Use File (SUF) of that wave. Up to now there are seven Scientific Use Files – one for each wave. Together they gather information on 44,854 valid interviews of 20,715 participants.

In 2011 (wave 4), a total of 4.854 individuals, aged 43-98 years, had participated at least twice. The Scientific Use File DEAS2011 with the data for the 2011 survey year contains information of participants from the following years of their first interview:

| Sample by baseline year | Number of respondents in wave 4 (2011) |
|---|---:|
| Panel respondents, B1996 | 1,039 |
| Panel respondents, B2002 | 957 |
| Panel respondents, B2008 | 2,858 |
| Total | 4,854 |

The sampling design of the German Ageing Survey allows for different perspectives of research, for example the option for cross sectional analysis of the individual living situations and the living conditions over a period of 21 years (see Klaus et al. 2017). Furthermore, the analysis documents social changes in a particular year of sampling over this period, starting from 1996 to 2017 (see Mahne et al. 2017), as well as estimating individual trends over three to 24 years with the panel data are possible (see Vogel et al. 2019). Another perspective arises from the comparison of individual change over six years in the time spans of 1996-2002, 2002-2008 and 2008-2014. This makes cohort analyses possible that studies the comparison of trends in constant age groups for different birth cohorts. One example is the prospective comparison of the transition from work to retirement or the health change in individuals for different cohorts.

For further information about the DEAS sample design, please visit Klaus & Engstler (2017), Engstler & Schmiade (2013), Engstler & Motel-Klingebiel (2010) and Engstler & Wurm (2006).

## 1.2 Topics and instruments of DEAS 2011

As in previous waves, respondents were asked in detail about their living situation. Information about the following subject areas was surveyed:

– Employment and retirement
– Partnership, family and intergenerational relationship
– Social networks and social support
– Leisure activities and voluntary work
– Housing situation and mobility
– Economic situation and economic behavior
– Subjective well-being
– Health and health behavior
– Need of assistance and need of care
– Attitudes, norms and values
– Basic sociodemographic data.

As in previous waves, the questionnaire is divided into two parts. The main part of information is collected by a face-to-face interview usually conducted in the respondent's home. The method of collecting data is a Computer Assisted Personal Interview (CAPI). In this oral interview the main large thematic fields of the DEAS are covered. A complex filtering makes a pinpointed group specific inquiry of differentiated information possible. Basically questions were posed as in the previous wave. Some questions could be omitted because they record permanent characteristics of the respondents, for example the highest degree of education or the childhood history. Additionally, a pulmonary function test to measure the lung capacity was conducted. Directly after the interview, the cognitive capacity of the participants was tested using a digit number test. In the second part, respondents were given a questionnaire to fill out by themselves ("drop-off"). In this part subjective beliefs – as opinions, values, images of old age, wellbeing - , more sensitive areas (for example details of the material situation and health) and topics do not need a differentiated filtering. In contrast to previous waves, more questions concerning the living environment were posed.

Further information on the design, topics, fieldwork and sample sizes of the 4th wave you will find in the section on methodology in the DEAS-report (available on request) as well as in the infas-methodological report (infas 2012). Instruments and methodological reports of all waves of the DEAS can be downloaded via the webpages of the <u>Research Data Centre (FDZ-DZA)</u>.

## 2   RESPONSE RATE AND LONGITUDINAL WEIGHTING

First inconsistency checks were undertaken by infas, followed by detailed data cleaning by the DZA. Interviews were deleted if there was suspicion that they were conducted via telephone or were not conducted at all or with the wrong target person. Finally 4,854 cases remain for analysis. The response rate is 58 percent of the adjusted gross sample, which is comparable to other longitudinal studies, analyzing the same age groups (Blom & Schröder 2011: 57).

*Data weighting*

In 2018 infas calculated completely new data weightings for the cross-sectional and longitudinal waves of the survey since 2002, replacing all previous weightings. This took place in close coordination with the DEAS-team. The chosen procedure correspondents to the familiar methods of all panel studies (for details, see Hammon et al. (2016), Trappman (2013), European Central Bank (2016)).

Basis of the longitudinal weighting were statistical default models (logistical regression) for the determination of the probability participation for the present and the future survey (excluded panel respondents, who died meanwhile), using a small range of Boolean predictors (see infas 2018: chapter 7.1).

To construct the longitudinal weights of the face-to-interviews the multiplying of the cross-sectional weighting of the first wave is calculated by the reciprocal of the participation probability in the next wave. For the weighting of the drop-offs statistical default models were

calculated, based on the participation in the face-to-face interviews. The longitudinal weighting for the drop-offs is calculated by multiplying the longitudinal weights for the face-to-face interviews by the reciprocal of the probability participation to the drop-offs. This method has been done for all waves since 2002, thus for each wave two longitudinal weighting variables (CAPI, drop-off) and two variables with the reciprocal of the probability participation probability exist (CAPI, drop-off). Combining the wave-specific probability participation longitudinal comparisons are possible (for details, see infas 2018: chapter 7.1). The Scientific Use Files of the survey waves do not include neither the longitudinal weighting variables nor the variables of the reciprocal values. The variables are retrieved in a separate file, which is only available for registered users.

Since 2002 an integrated cross-sectional weight has been done. To achieve this, each of the subsamples of a wave has been integrated to a collective cross-sectional sample. The subsamples represent (a) returnees ("Wiederteilnehmende"), (b) respondents with nonresponse in the wave before ("Rückkehrer nach temporärerem Ausfall) und (c) new respondents ("Erstbefragte"). Weighting for the integration is the real cross-sectional weighting of the base sample (see Engstler & Hameister 2016; chapter 2) and for returnees of the previous wave the longitudinal weighting of the present wave. For the returnees with temporary nonresponse two default models have been calculated on the base of auxiliary module: statistical default model and statistical model of return (for details see infas 2018: chapter 7.2). The weighting of the three subsamples have been combined by composite weighting ("Konvexkombination") to an integrated cross-sectional weighting ("Querschnittsgewicht") by multiplication the individual initial weight and the relative share of the subsample in the sample. Furthermore the probability to drop-off for all face-to-face-interviewees has been calculated on the base of a default model. The cross-sectional weighting of the drop-off is calculated by multiplication the cross-sectional weighting for the face-to-face-interview and the reciprocal participation probability of the drop-off.

The both integrated cross-sectional weightings of a wave are the basic distributions reveals in the micro-census of a particular year. The post-stratification is reported with iterative proportional fitting according to the indicators *age*gender*federal states, excluding the very small number of respondents, aged 91 and over. The Scientific Use File DEAS 2011 concludes exclusively the both weighting variables with the integrated post-stratification cross-sectional weightings for the face-to-face interview (qsps_11) and the drop-off (qspsdrop_11). The integrated cross-sectional weightings without post-stratification are retrieved in a separate file.[1]

Using the panel weights is in the responsibility of the individual user. The data weighting shows one of different possibilities for solving the problem of selective panel mortality. The cross-sectional representations is recommended to use descriptive presentation and statements. In the case of drop-of-respondents, please use the variable qspsdrop_11 (cross-sectional weightings), otherwise the variable qsps_11.

---

[1] The cross-sectional weighting for the DEAS wave 2002, 2008 and 2014 are only available for registered users by the Research Data Centre (FDZ-DZA).

# 3  DATA FORMATS

The Scientific Use File (SUF) 2011 as well as the SUF Meta (see chapter 5) are available in SPSS (version 22) and Stata (version 14) data format. If problems occur, please contact FDZ-DZA (fdz@dza.de) for advice.

# 4  CLASSIFICATION OF VARIABLE NAMES AND MISSING CODES IN THE DEAS 2011

In 2008 a new classification of naming of variables was introduced that is also used for the further waves. The variables were coded in the logic of a card scheme in the first two waves of the DEAS.[2] Since wave 3, the variables are coded on the basis of the question number. The question number is concurrently thematically grouped (for example all questions regarding the family situation are coded with 300), which makes it easier for the user to get along with the data as well as writing the analysis syntax.

The original variables in the survey year 2011 start with the character "g"[3] as wave marker. Variables from the CAPI interview are followed by the character "c". That means all variables from the oral interview start with "gc". Variables from the drop-off questionnaire are labeled with "gd". The following number is then the question number. For example, variable gc101 is question 101 from the CAPI questionnaire, variable gd5 if question 5 from the drop-off. Open item lists are coded with the ending "o" (gc106o). Multiple answers or answering options are differentiated using a serial numeric ending (gc423_1, gc423_2, etc.). Is there a fixed sequence of questions, for example as in the surveying on information for every child, there is a serial number at the end of every loop (gc3011, gc3012, etc.).

An entire overview of all variable names ever used in DEAS data can be found in the variable correspondence list on the web page on documentation of the FDZ-DZA. Basic values of all variables of the SUF DEAS 2011 including all labels and missing values can be found on the same web page in the codebook 2011.

The labeling of missing codes was basically altered in the DEAS wave 4 compared to previous waves. The new coding facilitates the automatic definition of missing values and harmonizes the reasons and specification for non-valid information. [4]

---

[2] For example: question 101 asking for the labour status (code "w") in the personal interview of wave 2 is placed on the second position at card 14 and is therefore labeled as w14_12.

[3] „g" was chosen because the drop-off variables from wave one are coded with „d", in 2002 with „e" and in 2008 with „f". The variables in the forthcoming waves will get a consecutive naming.

[4] Missing values in previous waves were coded as 0, -1, -2 or as 7, 8, 9 or 97, 98, 99 etc. respectively. A consistent definition of missing values for all variables in one data set was therefore not possible.

In the following table all codes for missing values are described:

| Value in SPSS | Value in Stata | Label [description] |
| --- | --- | --- |
| -1 | .a | Declined |
| -2 | .b | Don't know |
| -3 | .c | Filtered (question) |
| -4 | .d | Not asked (sample) [if question is either posed to panel respondents or basic sample respondents] |
| -5 | .e | No drop-off |
| -6 | .f | No answer [without closer differentiation] |
| -7 | .g | Deleted |
| -8 | .h | Double entry [if not deleted before] |

Generally the missing codes are formatted as missing labels. Before using them for deeper analyses, a frequency count is recommended.

## 5   META-DATA ON PARTICIPATION HISTORY, DROP-OUT REASONS AND MORTALITY OF DEAS RESPONDENTS

For easier using and overview the Research Data Centre (FDZ-DZA) provides sets of meta-data as an overview of the participation history of the respondents. The meta-data file contains information about all persons, who have ever participated in any of the survey questionnaires. It is easier to merge data sets and to analyse continuance. The following table shows all variables contained in the data set (date: January 2021):

| Variable | Label | Description |
| --- | --- | --- |
| fallnum | Case number (first digit for initial interview wave) | To merge information from all SUFs (1996-2020) |
| stich | Sample ID | Baseline sample of respondent |
| part_96 | Participation 1996 | Yes; no |
| part_02 | Participation 2002 | Yes – baseline sample 2002; yes – foreigner sample 2002; yes – panel; no |
| part_08 | Participation 2008 | Yes – baseline sample 2008; yes – panel; no |

| part_11 | Participation 2011 | Yes – panel; no |
|---|---|---|
| part_14 | Participation 2014 | Yes – baseline sample 2014; yes – panel; no |
| part_17 | Participation 2017 | Yes – panel; no |
| part_20 | Participation 2020 (compact survey) | Yes – panel; no |
| part | Participation 1996-2014 | Shows in which survey years respondent was interviewed |
| lastpart | Last participation | Year of last participation |
| n_obs | Number of participation | Number of individual panel participations |
| entry | Year of entry | Year of first interview |
| yob | Year of birth | Year of birth (possibly updated after plausibility check) |
| gender | Gender | Male; female |
| natdeutsch | German citizenship | Feature of the first interview: German, non-German |
| bildung4 | Level of education, 4-categories | Based on highest level of qualification in school and vocational college: low, medium, sophisticated, high |
| isced | Level of education, 3-categories (ISCED) | Low (ISCED 0-2), medium (ISCED 3-4), high (ISCED 5-6) |
| fiktiv_kreis_96 | Fictional number of district, 1996 | The real district numbers (since 31.12.2013) were replaced with a fictional number; this number can be used for multilevel analysis |
| fiktiv_kreis_02 | Fictional number of district, 2002 | Ditto. |
| fiktiv_kreis_08 | Fictional number of district, 2008 | Ditto. |
| fiktiv_kreis_11 | Fictional number of district, 2011 | Ditto. |
| fiktiv_kreis_14 | Fictional number of district, 2014 | Ditto. |
| fiktiv_kreis_17 | Fictional number of district, 2017 | Ditto |
| bbsr_kreistyp96 | Type of district, 1996 | Urban-rural typology (4 categories) |
| bbsr_kreistyp02 | Type of district, 2002 | Urban-rural typology (4 categories) |

| | | |
|---|---|---|
| bbsr_kreistyp08 | Type of district, 2008 | Urban-rural typology (4 categories) |
| bbsr_kreistyp11 | Type of district, 2011 | Urban-rural typology (4 categories) |
| bbsr_kreistyp14 | Type of district, 2014 | Urban-rural typology (4 categories) |
| bbsr_kreistyp17 | Type of district, 2017 | Urban-rural typology (4 categories) |
| bbsr_kreistyp20 | Type of district, 2020 | Urban-rural typology (4 categories) |
| rlc2002_kat | Return code 2002 | Reasons for not participating in wave 2 (or code for participation) |
| rlc2008_kat | Return code 2008 | Reasons for not participating in wave 3 (or code for participation) |
| rlc2011_kat | Return code 2011 | Reasons for not participating in wave 4 (or code for participation) |
| rlc2014_kat | Return code 2014 | Reasons for not participating in wave 5 (or code for participation) |
| rlc2017_kat | Return code 2017 | Reasons for not participating in wave 6 (or code for participation) |
| vitalstatus_last | Last known vital status | Notes if person is dead (0) or alive (1) |
| vitalstatus_source | Year of last information on vital status | Either year of survey field work or year of address update (including information form registration offices) |
| surv_birth | Survival period in months | Months since birth until death or end of observation |
| surv1996 | Survival period since interview 1996 | Months since interview 1996 until death or end of observation |
| surv2002 | Survival period since interview 2002 | Months since interview 2002 until death or end of observation |
| surv2008 | Survival period since interview 2008 | Months since interview 2008 until death or end of observation |
| surv2011 | Survival period since interview 2011 | Months since interview 2011 until death or end of observation |
| surv2014 | Survival period since interview 2014 | Months since interview 2014 until death or end of observation |

Especially the development of a longitudinal data set is facilitated by using the Meta dataset. Information on drop-outs and on participation in panel questionnaires opens up the possibility of mortality and drop-out analyses. Data users who want to analyse reasons of panel attrition may contact the Research Data Centre (FDZ-DZA) to get more details about return codes.

## 5.1 Profile of participation

The variable *part* combines information on all seven DEAS survey years. All respondents get a 6 digit value: 1111111 for example shows that this person participated in all seven DEAS waves, respondents with the number 0100111 were firstly asked in 2002, did not participate in 2008 and 2011, but answered the questionnaire again in 2014, 2017 and 2020.

The sample to be analysed can be easily compiled with the help of variable *part*. If one wants to conduct a longitudinal analysis for example, respondents can be chosen on the basis of the participation profile (variable *n_obs*). The personal information from the SUFs can be added to the selected cases with the help of the personal identifier *fallnum*. If you have any questions in using this Meta dataset please consult the FDZ: [fdz@dza.de](mailto:fdz@dza.de)

## 5.2 Data on survival of respondents

All respondents were contacted by the field research institute before the next panel questionnaire as well as in panel maintenance or in a special survey of whereabouts of previous respondents. This was done to record the recent vital status. In case of death the exact date of death was detected by the registration office. The variable *vitalstatus_last* indicates the last known vital status of all DEAS respondents (0=alive, 1=dead). Respondents that were only interviewed once and were not willing to participate again, maintain the status they had at the first interview. Otherwise the last information booth is used. Variable *vitalstatus_source* contains the calendar year from which we know the last known vital status. Variable *surv_birth* contains the number of months since birth until death or last known survival status. Variables *surv_birth, surv1996, surv2002, surv2008, surv2011, surv2014* and *surv2017* (survival since the respective date of interview) – together with the variable *vitalstatus_last* (death or censoring indicator) - enable to do survival analysis of the risk of death.

*Variables rlc2002_kat* to *ric2017_kat* give the return codes of all respondents, who were selected to be part of the panel samples for the DEAS surveys in 2002 up to 2017. The return codes are categorized and provide information if the target person could be interviewed or not. If there was no interview possible, reasons for that are listed. Interviews that were deleted later in the process of data cleaning get code 6.

## 6 CHANGES IN THE DEAS 2011 AS IN CONTRAST TO THE DEAS 2008

## 6.1 New questions

There are only a few new questions in the DEAS survey of 2011 – none in the CAPI. Following questions and items are new in the drop-off questionnaire:

| Variable | Label |
|---|---|
| gd55_8 | Household equipment: MP3-Player |
| gd55_9 | Household equipment: game console |
| gd55_11 | Household equipment: internet |
| gd55_14 | Household equipment: sports equipment or home trainer |
| gd56_1 | Freedom of barrier: accessible without steps |
| gd56_2 | Freedom of barrier: staircase has handrail on both sides |
| gd56_3 | Freedom of barrier: rooms have no or low doorsills |
| gd56_4 | Freedom of barrier: bathroom and toilet have wide doors, accessible to wheelchairs |
| gd56_5 | Freedom of barrier: door of bathroom opens outwards |
| gd56_6 | Freedom of barrier: all rooms accessible without steps |
| gd56_7 | Freedom of barrier: bathtub exists |
| gd56_8 | Freedom of barrier: barrier-free shower exists |
| gd56_9 | Freedom of barrier: non barrier-free exists |
| gd57_1 | Using of new technical development: individually programmable control |
| gd57_2 | Using of new technical development: emergency call system |
| gd57_3 | Using of new technical development: household appliances with special safety function |
| gd57_4 | Using of new technical development: other |
| gd57_4o | Using of new technical development: other, open question |
| gd57_5 | Using of new technical development: no, not used |
| gd58_2 | Neighbourhood: barrier-free adjustments should be made |
| gd58_10 | Neighbourhood: correspondents to own ideas and needs |
| gd58_11 | Neighbourhood: there are residential and nursing homes |
| gd58_12 | Neighbourhood: like to spend free time here |
| gd58_13 | Neighbourhood: live people I can rely on |
| gd58_14 | Neighbourhood: many possibilities for walking |
| gd58_15 | Neighbourhood: many possibilities to do sports |
| gd58_16 | Neighbourhood: enough seats are availaible |
| gd58_17 | Neighbourhood: accessible public toilets |

## 6.2 Changes in filtering and other omitted questions

In contrast to the CAPI in 2008 there are changes in filtering of panel participants. This was mainly done to reduce the time burden for respondents in skipping some additional

questions. This filtering was mainly done for characteristics that are assumed to be stable since previous waves. Before preparing the data set for analysis one should check if missing are due to filtering, data revision or answering behavior of respondents. For this purpose, the new detailed missing scheme is especially helpful. If time invariant characteristics are filtered at one point of time, they can be taken from SUFs of previous waves.

There is a relevant change in the questions regarding volunteering: in the DEAS 2011 the question 416 concerning volunteering outside of groups, associations or organizations was only posed to respondents, who are no member at all. Question 416 in 2011 is again posed to all respondents as in 1996 and 2002 and is hence not as restrictive as in 2008.

Question 434 (general images of ageing) in the CAPI is dropped as well as question 8 (political participation), 9 (relationship of older to younger people), 10 and 11 (area specific government purchases and responsibility), 45 (health care proxy and patient's provision), 54 and 55 (values), 56 (sexual orientation) and 75 (preferences of the distribution of money) in the drop-off. General questions on attitudes on responsibility for the provision in old-age are transferred from CAPI to the drop-off questionnaire.

## 6.3   Changes on phrasing of questions

There are only minor changes made in the phrasing of questions in the DEAS 2011 in comparison to former waves. With the help of variable labels and questionnaires we recommend users to check if questions or items have changed when comparing variables over waves.

## 7   APPLIED MEASURES TO ANONYMIZE DATA IN THE SUF

To ensure the factual anonymity of respondents, some variables are not included in the SUF DEAS2011. Other variables are oversimplified in some characteristics.

− Some variables that contain the regional context are deleted because in some cases it might be possible to determine the living area of the respondent. Also deleted are information on birthday and month of birth of respondents.
− Some variables with open answer option are deleted or specific answers were generalized. Such changes at open answers were marked with the sign #.
− Rare demographic characteristics of family formations were oversimplified. The few people with the attribute "civil union" are assigned to the category "married, living together with spouse".
− Rare year specifications were merged to categories. This was done by generating a new categorical variable (identifiable through the ending "kat" of the variable name) or by recoding this value in an existing numerical variable to the next value (with the appropriate designation of the value labels).
− Partly maximum values are categorized (for example number of siblings, household size).

There are secured PC workplaces available for guest researchers within the DZA. Please contact the consulting of the FDZ (fdz@dza.de) if you want to conduct research using the full DEAS data to do research on the level of districts or including some context characteristics of the immediate neighbourhood for example.

# 8 GENERATED VARIABLES

There are various generated variables added to the Scientific Use File (SUF) DEAS 2011. This helps to compare constructs over time and with other data sources and simplifies the entry into the data analysis. If you need the syntax files for the generated variables, please contact the FDZ: fdz@dza.de

Overview on generated variables in the SUF DEAS2011 (Status: May, 2019)

| Variable | Label | Available for the years | | | |
|---|---|---|---|---|---|
| | | 1996 | 2002 | 2008 | 2014 |
| fallnum | Case number (first digit for initial interview wave) | x | x | x | x |
| stich | Sample idenfication | x | x | x | x |
| qsps_11 | Cross-sectional weight – baseline sample oral interview | x[1] | x | x | x |
| qspsdrop_11 | Cross-sectional weight – baseline sample drop-off | x[1] | x | x | x |
| part_96 | Participation 1996 | x | x | x | x |
| part_02 | Participation 2002 | x | x | x | x |
| part_08 | Participation 2008 | x | x | x | x |
| natdeutsch_11 | German nationality[3] | x[1] | x | x | x |
| migrat_11 | Migratory background[3] | | x | x | x |
| westost_11 | Western or eastern part of Germany | x | x | x | x |
| bland_11 | Federal state | x | x | x | x |
| polgk_11 | Community size | x | x | x | x |
| bikgk10_11 | BIK region size | x | x | x | x |
| bbsr_kreistyp_11 | Type of district | | | | |
| alter_11 | Age (year of interview minus year of birth) | x | x | x | x |
| altervoll_11 | Consummate age at interview | x | x | x | x |
| bildung4_11 | Level of education, 4 categories[3] | x | x | x | x |
| isced_11 | Level of education by ISCED, 3 categories[3] | x | x | x | x |
| erw_11 | Labour force participation 2011 | x | x | x | x |
| isco88zp3_11 | ISCO88-code of (last) occupation of respondent | (x)[2] | x | x | (x)[2] |
| isco88comzp3_11 | ISCO88com-code of (last) occupation of respondent | | x | x | |

| Variable | Label | Available for the years | | | |
|---|---|---|---|---|---|
| | | 1996 | 2002 | 2008 | 2014 |
| isco88partner3_11 | ISCO88-code of (last) occupation of present or last partner | (x)[2] | x | x | (x)[2] |
| isco88compartner3_11 | ISCO88com-code of (last) occupation of present or last partner | | x | x | x |
| siops_11 | SIOPS – occupational prestige – based on couple (Treiman) | x | x | x | x |
| siops_kat_11 | SIOPS – occupational prestige, categorized | x | x | x | x |
| mps_11 | MPS occupational prestige – based on couple (Wegener) | x | x | x | |
| isei_11 | ISEI occupational status score – based on couple (Ganzeboom, De Graaf & Treiman) | x | x | x | x |
| schicht_11 | Social strata – based on couple (Mayer & Wagner) | x | x | x | x |
| egpzp_11 | Social class (EGP) of respondent (Erikson, Goldthorpe & Portocarero) | | | x | |
| egppartner_11 | Social class (EGP) of (last) partner | | | x | |
| eseczp_11 | Social class (ESeC) of respondent (Rose & Harrison) | | | x | |
| esecpartner_11 | Social class (ESeC) of (last) partner (Rose & Harrison) | | | x | |
| hheink_11 | Monthly net household income, from CAPI and drop-off | x | x | x | x |
| aee_oecd11 | Monthly equivalence income (new OECD scale) | x | x | x | x |
| einkpos_11 | Income position (% of mean of equivalised income of population) | x | x | x | x |
| einkarm_11 | Income poverty (< 60 % of median equivalent income of population) | x | x | x | x |
| einkreich_11 | Income wealth (>200 % of mean equivalent income) | x | x | x | x |
| partner_11 | Existence and type of partnership | x | x | x | x |
| exkind_11 | Existence of living children | x | x | x | x |
| anzkind_11 | Number of living children | x | x | x | x |
| wekind_11 | Geographical distance to nearest living child | x | x | x | x |
| exenk_11 | Existence of living grandchildren | x | x | x | x |
| anzenk_11 | Number of living grandchildren | x | x | x | x |
| enkelbetreu_11 | Care of grandchildren | x | x | x | x |
| nwgroesse_11 | Size of personal network: number of important people in regular contact | x | x | x | x |
| ehramt_11 | Voluntary work in groups and organizations | x | x | x | x |
| ehramt_weit_11 | Voluntary work in general | x | x | x | x |
| bmi_11 | Body-Mass-Index | | x | x | x |

| Variable | Label | Available for the years | | | |
|---|---|---|---|---|---|
| | | 1996 | 2002 | 2008 | 2014 |
| sf36_11 | Physical Functioning Scale (SF 36 Short Form) | | x | x | x |
| anzphy_11 | Total number of physical diseases | x | x | x | x |
| lft_11 | Pulmonary function test (peak flow in l/min) | | x | x | x |
| zzbearb_11 | Handling of the digit symbol test (ZZT) | | x | x | x |
| zzgscore_11 | ZZT: total score of recorded digits | | x | x | x |
| zzfscore_11 | ZZT: total score of wrong recorded digits | | x | x | x |
| zzrscore_11 | ZZT: total score of correct recorded digits | | x | x | x |
| depressiv_11 | Depression Scale CES-D (Hautzinger) | | x | x | x |
| lone6_11 | Scale for Loneliness (De Jong & Gierveld) | x | x | x | x |
| selbstwert_11 | Self Esteem Scale (Rosenberg) | | x | x | x |
| optimismus_11 | Optimism Scale (Brandtstädter & Wentura) | | x | x | x |
| selbstwirk_11 | Generalized Self-Efficacy Scale (Schwarzer & Jerusalem) | | | x | x |
| hope_11 | HOPE Scale (Snyder et al.) | x | x | x | |
| flex_11 | Flexible Goal Adjustment Scale (Brandtstädter & Renner) | | x | x | |
| lz_11 | Scale of Life satisfaction (Pavot & Diener) | x | x | x | x |
| pa_11 | Positive affect (Watson, Clark & Tellegen) | x | x | x | x |
| na_11 | Negative affect (Watson, Clark & Tellegen) | x | x | x | x |

1) With different variable names due to different measurement.
2) For 1996 the ISCO codes are only in the ISCO68 classification, for 2014 only in the ISCO08 classification.
3) Information given in the starter interview.

The listed variables are described shortly in the following:

## 8.1 Identification number

The variable *fallnum* contains a respondent's code number, which differs from the original ID of the interviewed person. It is a seven-digit number: the first three digits show the wave on which the person participated the first time (B1996: 100, B2002 and A2002: 200, B2008: 300, B2014: 400). The last four digits are the actual respondent's code numbers. All respondents keep their code number from the first survey for all upcoming surveys. By using the ID-variable *fallnum* the different datasets of a person can be combined longitudinally.

## 8.2  Sample population

The variable *stich* marks the sample's origin and the year of the first interview. Possible values up to 2011 are 1 to 3 (1 = person origins in baseline survey 1996; 2 = from baseline 2002; 3 = from baseline 2008; 4 = from baseline 2014).

## 8.3  Weighting

The variable *qsps_11* contains the post-stratification cross-sectional weighting for all interviewees up to the age of 90. The weighting is normed by the number of cases (arithmetic mean=1). The variable *qspsdrop_11* contains the cross-sectional weighting for the drop-off information of this baseline sample. The weighting is normed by the number of cases (arithmetic mean=1). If the countings are limited to drop-off participants, the variable *qspsdrop_11* is recommended, otherwise the variable *qsps_11*. The use of cross-sectional weightings are provides for descriptive analysis.

More information on constructing the panel weights is available in chapter 2.

## 8.4  Longitudinal participation

The variables *part_96*, *part_02* and *part_08* show if a respondent of the year 2011 has also participated in an interview in 1996, 2002 and 2008 and if the participation in a previous wave has been the first interview (code 1) or a follow-up survey (code 2).

## 8.5  Nationality

The variable *natdeutsch_11* informs about the nationality of interviewees of the baseline sample. It shows whether the person owns the German nationality. The information dated back to from the self-reported data of the interviewees to their nationality, which is recorded only for the baseline sample. The nationality of the panel participants, if required, is to be taken from the first measuring time. Up to baseline 2008 information about the citizenship were given by the registration office.

## 8.6  Migration background

The DEAS-generated variable *migrat_11* covers the migration background of the interviewees of the baseline sample. It follows the concept the Federal Statistical Office uses in the microcensus. Three different groups are distinguished: People without migration background, people with migration background (i.e. migration to Germany) as well as people with migration background but without an own migration experience (i.e. born and grew up in Germany). Therefore the details of the first interview about nationality, place of birth, year of immigration, possession and year of acquisition of the German or a foreign nationality was

used. Immigration before 1950 does not count as migration background. Those, who are born in the former eastern regions or the 'German Reich', immigrating to the FRG or GDR after 1949 are counted as people with migration background.

## 8.7   East-West allocation and Federal States

Based on the current residential address infas carried out an assignment to the region of the former federal territory (before German re-unification, including the western parts of Berlin) or the former East Germany (including the eastern part of Berlin) in variable *westost_11.*

The variable *bland_11* shows in which federal state the respondent's place of residence is located.

## 8.8   Size of municipality and region

The variable *polgk_11* contains the categorized population of the respondent's municipality of residence. The variable *bikgk10_11*, also created by infas, includes the categorized population of the BIK-region to which the municipality of residence belongs. The BIK-regions are an area of interdependence which show the relationships between cities and their surrounding areas for conurbations, middle- and local sub-centres. For more information visit: https://www.bik-gmbh.de/cms/regionaldaten/bik-regionen

## 8.9   District types

Variable *bbsr_kreistyp_11* contains the settlement information of the district type of the country or the independent city in which the interviewee lives. Basis is the new typification of the different structural characteristics of settlements conducted by the Federal Institute for Research on Building, Urban Affairs and Spatial Development into four district types (BBSR 2012).

## 8.10  Age

The variable *alter_11* contains the age reached in the year of data collection as a difference between year of birth and year of the survey. The variable *altervoll_11* derived from the date of birth indicates the completed year of age on the day of interview.

## 8.11  Education

The Scientific Use File contains two variables to the educational level. Based on information to the general and professional education at school and vocational training, especially for the achieved level, a 4-stage variable about the educational level are offered (*bildung4_11*). In

addition, a 3-stage educational construct is made (*isced_11*), referring to the ISCED categories (Internationally Standard Classification of Education). Kind of educational information collected depends on the place of education (in Germany or abroad).

*4-stage level of education*

The allocation to variable *bildung4_11* occurs following this systematic:

| | |
|---|---|
| Level 1 (low) | Respondents without completed vocational qualification and up to a maximum of a graduation degree, which qualifies for a professional qualification |
| Level 2 (medium) | Respondents with vocational qualifications or qualifications for university or university of applied science entrance |
| Level 3 (sophisticated) | Respondents with finished upgrading training (professional, master or technical school, university of cooperative educations or academies) |
| Level 4 (high) | Respondents with completed university studies (university or university of applied science) |

*Level of education according to the* ISCED scale

The allocation to variable *isced_11* occurs following this systematic:

| | |
|---|---|
| Level 1 (low) | ISCED 0-2; respondents without formal vocational qualification |
| Level 2 (medium) | ISCED 3-4; respondents with vocational training (at work or at school), including respondents with higher general school certificate without professional training |
| Level 3 (high) | ISCED 5-6; respondents with completed professional development training (professional, master or technical school, university of cooperative educations or academies) and respondents with completed university studies (university or university of applied science) |

## 8.13 Occupational status

In the variable *erw_11* three groups are distinguished:

−   Employed persons; these are persons, who answer in question 101 that they are employed full-time or part-time or irregular, marginally employed or working in a secondary job.
−   In retirement; these are persons aged 60 or older, who receive own old-age pension (question 100) or who are in early retirement (question 101). The classification to this category took place regardless of a possible employment during the retirement, as *erw_11* questions the main status.
−   Others not employed persons; these are persons with other valid codes in question 101 and persons never working full-time in accordance with question 32a. Also employees in the release phase of early retirement with zero working hours (question 101) are added to this category.

## 8.14 Classification of occupation, prestige and status

### 8.14.1 ISCO88-Codes

On the basis of the open answers to the actual or last occupational activity the ISCO88-codes (International Standard Classification of Occupations) and ISCO88com-codes were allocated[5]. These three-digit codes replace the open answers and are positioned in the appropriate interview parts (e.g. *gc132isco88_3*). ISCO codes exist for the respondent as well as for his/her current partner or the last spouse.

If there was no change of the occupational situation since the last interview (e.g. constantly retired), the open question asking for the exact occupation was filtered. The same is true for the occupation status of the current partner or - for divorced persons without a partner - the former spouse. For constructing variables *isco88zp3_11* and *isco88comzp3_11* (current of former occupation of the interviewee) as well as variables *isco88partner3_11* and *isco88compartner3_11* (occupation of partner or former spouse) information from previous waves was used if required.

### 8.14.2 SIOPS-occupational prestige

The „SIOPS - Standard International Occupation Prestige-Scale" was proposed by Treiman (1977) and further developed by Ganzeboom & Treiman (1996, 2003). Every single ISCO88-code is allocated a specific value of prestige, that measures mean social reputation of that occupation in the society. The scale is based on research to evaluation the occupations in the society in various countries. The scale lasts from 13 to 78 in the DEAS survey data.

---

[5] The ISCO-coding to the actual or last occupational activity (wave 4) and the providing of the single variables for the professional prestige (SIOPS, MPS), the socio-economic status (ISEI) and the social class (EGP, ESeC) have been carried out by infas, Institute for Applied Social Sciences, under the authority of the DZA (German Centre of Gerontology). There are only minimal differences in the ISCO88com- and ISCO88-codes.

On the basis of the ISCO88-codes of the current or last occupation, a value of prestige was allotted to every target person. This was also done for the current partner or – for divorced, separated or widowed persons without partner – the last spouse. Relevant was the current or last occupation of the partner during the marriage. If there was no change in the occupation since the last interview, the in-depth occupational activity questions were filtered and information from previous waves were taken to generate the index.

Following, both values of prestige were transformed into one single household or partner-related level of prestige. In variable *siops_11* the higher value of both persons were taken. That means, if the present or the former partner (for persons without a partner) has a higher value of prestige, this was taken to generate the variable.

In addition to variable *siops_11* with the differentiated metrical scale values the variable *siops_11_kat* was derived. The scale values were collapsed to five categories from 1 (low) to 5 (high) (see Hoffmeyer-Zlotnik 2003).

### 8.14.3 MPS-occupational prestige

The Magnitude Prestige Scale (MPS) is as well based on the ISCO scale and constructed by Bernd Wegener (1988) for Germany. Variable *mps_11* has a valuation from 20.0 to 186.8 points of prestige and is also based on the couple.

### 8.14.4 ISEI-Status-Score

The „ISEI – International Socio-Economic Index of Occupational Status" is a scale constructed by Ganzeboom et al. (1992) to measure socio-economic status. It is based on the assumption that every occupation needs a mean educational background and enables a specific level of income. Basis for the construction of the scale was a study conducted in 16 different countries. Educational as well as income values for various occupational representatives were identified. Every occupation within the ISCO classification is allotted a status score.

On the basis of the ISCO88-codes of the current or last occupation a status score was identified for every target person. This was also done for the current partner or last spouse if the interviewee is separated, divorced or widowed. The current or last occupation of the partner respectively the last occupation during the divorced marriage with the former spouse was decisive for the construction of the partner variable. In case there was no change in the occupation in 2011 since the last interview, information for panel participants from previous waves were taken to fill missing values.

Finally, variable *isei_11* is constructed using information on the target person as well as from the partner. The higher value from both partners was taken as household or partner related level of status. That means if the status score of the current partner or former spouse in the case of interviewees without a partner is higher than the score of the target person, the value of the partner is used in constructing *isei_11*.

### 8.15 Social strata and class

#### 8.15.1 EGP-class scheme

The EGP class scheme (Erikson et al. 1979; Erikson & Goldthorpe 1992) labels the social class of persons or household according to their position on the labor market, based on occupation and employment status. The main differentiating element is between employers, self-employed (with or without employees) and employees as well as the kind of regulation of the employment (classical labor control, employment relationship, mixed type). The employment relationship respects the degree of labor autonomy as well as the different forms of scrutiny and the empowerment to control. In contrast to hierarchical concepts of prestige or status, the EGP class scheme is a categorical scheme of differentiation. Infas executed the classification on behalf of the DZA and followed the operationalisation and naming of Ganzeboom & Treiman (2003).

Variable *egpzp_11* is based on different characteristics of the current or last occupation of the interviewee. It contains the categorical class of the target person. This was equally done for the current partner or the last spouse and stored in variable *egppartner_11.* In case there was no change in the occupation in 2011 since the last interview, information for panel participants from previous waves were taken to fill missing values. There is no partner related EGP variable as there is no singular hierarchy of class.

#### 8.15.2 ESeC-class scheme

The ‚European Socio-economic Classification' is a further development of the EGP scheme. It was developed by David Rose and Eric Harrison and commissioned by Eurostat to harmonize the European statistics for comparative analysis in the field of social inequality (Rose & Harrison 2010). Main basis for the classification is again the kind of regulation of the employment in combination with the job-related human capital and the performance monitoring. There are nine different ESeC classes. They are operationalised using the ISCO codes, the occupational status, the number of employees of self-employed and the oversight capacity.

Variable *eseczp_11* contains the class of the interviewee based on characteristics of the present or last employment. This is equally done for the partner or the former spouse and stored in variable *esecpartner_11.* In case there was no change in the occupation in 2011 since the last interview, information for panel participants from previous waves were taken to fill missing values. The official English terminologies are used (see Rose et al. 2010). A German translation of the categories is done by Wirth & Fischer (2008).

#### 8.15.3 Social strata

The partner related social strata *schicht_11* was constructed using the information on current or last occupation of the target person and the current partner (see Mayer & Wagner (1999) and Kohli (2000)). If widowed or divorced interviewees have no partner, information of the last spouse was taken to construct the variables. Assigned is the highest strata of the couple. According to the following table, five different occupational positions are differentiated in the variable *schicht_11.*

| Social strata: | Codes of the occupation: |
|---|---|
| Lower class | 10, 11, 60, 61, 62 |
| Lower middle class | 12, 25, 40, 41, 50, 51, 63, 64 |
| Middle class | 13, 14, 30, 35, 36, 42, 52, 55, 65 |
| Upper middle class | 31, 32, 43, 53 |
| Upper class | 20, 21, 22, 23, 33, 34, 44, 54 |

As in Mayer & Wagner (1999) we avoided to label the highest category as upper class because the societal elite (top manager, top-ranking politicians, large scale manufacturer) are only limited represented in the sample. In case there was no change in the occupation in 2011 since the last interview, information for panel participants from previous waves was taken to fill missing values.

## 8.16 Income

### 8.16.1 Monthly net household income

The income of target person and household is queried both in interview and drop-off. During the interview target persons are asked to provide the household net income (question 802). If they are unwilling or unable to do so they can choose from a list of income categories (question 803). Detailed information about the individual sources of income is gathered for both the target persons and their partners within the drop-off questionnaire (questions 68 to 70). In case of significant deviation between the interview's monthly income statement and the drop-off questionnaire's summed income, the data was related the various other characteristics of the target person and household, and adjusted accordingly for plausibility. Therefore, the final data set includes the new variable *gc802neu*. It contains either the actual answer to question 802 or, if unavailable, the income category's mean average of question 803 or, if unavailable or not plausible, no value.

Extending *gc802neu* the data includes variable *hheink_11* as part of the amended generated variables. It contains the household net income of variable *gc802neu* if available; else it contains the summed income based on the answers from the drop-off questionnaire. However, this is suppressed if the drop-off questionnaire states further household income - additional to the target person's and his partner's- of unspecified height (question 69).

### 8.16.2 Equivalent income

Variable *aee_oecd_11* contains the needs-adjusted monthly per head income of the household members of the target person. Criterion of the weighting is the so called modified OECD equivalent scale that is used by Eurostat and the German Federal Statistical Office. In this scale the first or single household member - older than or 15 years old - gets the weighting factor 1.0. Any other household members from the age of 15 get the factor 0.5 and

household members younger than 15 years get the factor 0.3. The sum of these weighting factors is the divisor for calculating net household income into equivalent income. Monthly household net income is taken from variable *hheink_11*.

### 8.16.3 Income position, income poverty and income wealth

Variable *einkpos_11* shows the individual income position in percentage points of the mean equivalent income of the whole population. Mean and weighted equivalent income from the GSOEP for the year 2011 amounting of 1.599.90 Euro serves as reference value.[6]

The dummy variable *einkarm_11* provides information if the equivalent income is below the poverty line. 60 percent of the median of the equivalent income is mainly used as poverty line In the German and European social statistic. Following this definition, the poverty line for the year 2011 is 840 Euro of the equivalent income on the Basis of GSOEP data.

Dummy variable *einkreich_11* provides information about income wealth. We use the same threshold that is used in the Poverty and Wealth Report of the Federal Government: having more than twice of the mean equivalent household net income is considered being wealthy. Using the GSOEP, the monthly equivalent income threshold is 3.200 Euro in the year 2011.

## 8.17 Partnership

Variable *partner_11* informs about the typo of household and partnership. It is differentiated between persons without a partner, persons with a partner in the same household and persons with a partner living outside the targets' household (living apart together).

## 8.18 Existence and number of children and grandchildren

Information on the existence and number of currently living children or grandchildren of the target person are stored in variables *exkind_11*, *anzkind_11* and *exenk_11* and *anzenk_11*. Information on children 1 to 8 form the basis of the calculations.

Interviewees telling the interviewer in question 300 about all children, who grew up or still grow up with them. Persons said that they have no children but mention children later on in the interview (for example by listing them in questions regarding the household composition or the support of family members) are considered as having children in variable *exkind_11*. Persons with children that are already dead, are considered as persons without children (code 0). Persons with children but refusing to give more information if the child is still alive are set on the value -6.

Variable *anzkind_11* is the sum of all living children. Persons with children, but refusing to give more information if the child is still alive, are set on the value -6. Persons with more than

---

[6] We thank Peter Krause, German Institute of Economic Research Berlin for providing the reference values for calculating income position, poverty and income wealth.

8 children are only asked for more information of the first 8 children. There is no information from the ninth child on. The maximum value of variable *anzkind_11* is therefore 8.

The proceeding for the existence (*exenk_11*) and the number of grandchildren (*anzenk_11*) is alike that for the children. The only difference is that only living grandchildren are taken into account.

## 8.19 Distance to the closest child

Variable *wekind_11* contains information about the distance to the closest living child. This variables has seven categories, ranging from „living in the same household" to „living abroad". Persons without children are filtered.

## 8.20 Caring for grandchildren

The construct *enkelbetreu_11* contains only information about persons with grandchildren. Grandparents who care for one or more grandchild get the code 1, persons having grandchildren but do not participate in caring for them get the code 0. Persons without grandchildren are filtered.

## 8.21 Size of network

Questions 600 to 607 in the oral interview serve as basis for constructing the network size. Variable *nwgroesse_11* contains the number of persons that are named as important persons to the target person with regular contact. If the target persons wanted to name more than 8 persons (question 607), the network size was set to 9+.

The approach taken here is just one possibility to measure network size. Often persons that are mentioned in the children and household matrix, are not named again. This gives a hint that variable *nwgroesse_11* only measures the minimum value of important persons with regular contact.

## 8.22 Volunteering, honorary posts

On the basis of questions 408 and 414 from the oral interview, the variable *ehramt_11* specifies if a person executes an honorary office in the groups or organizations in which he is a member. Other voluntary offices outside these groups and organizations (question 416) are included in variable *ehramt_weit_11*.

## 8.23 Body-Mass-Index

The body mass index is calculated as division of body weight (in kg) and square of body height (in meters). The unit of the formula is therefore kg/m$^2$. The variable *bmi_11* is a rough indicator for the evaluation of the weight of a person. For interpretation age as well as sex (and typically also amputated extremities) should be taken into account. Generally the normal weight calculated as BMI for men is between 20 and 25 kg/m$^2$. For women normal BMI weight is between 19 and 24 kg/m$^2$. A BMI value lower than 16 indicates heavy underweight whereas a BMI higher than 40 indicates adiposity of the third degree.

There are no general rules in handling extreme BMI values. In research potentially implausible values are tested in imitation of the trimming of values at García Villar & Quintana-Domeque (2009): when generating the BMI construct an analysis of plausibility of all BMI values lower than 15 and higher than 45 was undertaken. A comparison with the survey data in 2008 shows that all persons with noticeable values also had these extreme values three years before. There is only one exception that is set on -7 (implausible). Responsibility for the wise handling of the extreme values of the BMI is left to the researcher.

## 8.24 Physical functioning

The SF-36 subscale (Bullinger & Kirchberger 1998) is used to measure physical functioning. The degree of physical impairment is measured using an evaluation of ten daily activities on a scale from 1 (yes, limited a lot) to 3 (no, not limited at all). The sum of the items is then transferred into the standard 0-100 range. Higher values of variable *sf36_11* indicate a better physical functioning.

## 8.25 Number of physical diseases

Variable *anzphy_11* contains the number of physical diseases on the basis of question 32 of the drop-off questionnaire.

## 8.26 Pulmonary function test

The pulmonary function test is based on the peak flow method (recording of the maximum breath out capacity). The peak flow measurement represents a useful approximation of the vital capacity of people. Variable *lft_11* is consistent with the maximum value of two measurements during the interview. If there was only one measurement this value was taken for constructing the variable. The method is in accordance with the standard method described at Nunn & Gregg (1989) (besides the measurement of two instead of three values). The data corresponds to liter of exhaled air per minute. The measurement is accurate to +/- 10l/min when conducted correctly (according to manufacturers' instructions).

## 8.27 Digit Symbol Substitution Test

In accordance to the "Digit Symbol Substitution Test" (Wechsler 1955; Tewes 1994) used in intelligence tests, a digit symbol test is carried out with all interviewees. The interviewee is shown a table with codes of Arabic figures 1-9 that correspond to simple geometric signs. Then they have 90 seconds to fill out a table with four rows of figures with the corresponding geometric sign. The number of figures exceeds the number of possible entries in the given time. The interviewer notes problems while conducting the test. The digit symbol test is easy to execute and is therefore used more often in surveys (Hoyer et al. 2004: 211).

There are four generated variables that contain information on process and results of the digit symbol test. Variable *zzbearb_11* indicates if the test was conducted at all and if yes, is there a problem noted. Very high numbers of correct entries that are impossible to achieve as well as interviews with problems recorded by the interviewer are excluded and counted as invalid. The total number of characters is contained in variable *zzgscore_11*. The number of wrong characters is stored in variable *zzfscore_11* and finally *zzrscore_11* contains the deviation of all mentioned and wrong characters and therefore specifies all correct characters.

The test values are a good measurement for the cognitive performance of adults[7]: typically correlations of about -.46 to -.77 between age and test result are reported (Hoyer et al. 2004: 211).

## 8.28 Depressiveness

We use the short form of the German translation of the CES-D (Center for Epidemiologic Studies Depression) Scale (15 items, questions 504 of the interview) (Hautzinger & Bailer, 1993). It measures the depressive symptoms. The value of the scale is the sum over 15 items. First, a mean score is generated when at least 8 items have valid values. This mean score is then multiplied by 15 to get the sum score with a value range 0 to 45. High values of variable *depressiv_14* indicate higher depressive symptoms

## 8.29 Loneliness

We use the short version with six items (question 30 of the drop-off; the statement "I prefer being with people in my own age group" is not part of the scale) of the original scale to measure loneliness (De Jong Gierveld & Van Tilburg 2006). The value of the scale is the mean value the six items. At least three of the four items must contain valid values to calculate the mean. High values of variable *lone6_11* indicate a high level of loneliness.

---

[7] The test result is a global indicator of the cognitive performance that measures the speed of visual perception and information processing on the one hand. But besides that it also measures the pace in which the results are either written down or typed into the computer.

## 8.30 Self-Esteem

The scale *selbstwert_11* measures self-esteem (Rosenberg 1965). To calculate it the following ten drop-off items were used: question 2, items 1, 3, 4 and 6; question 16, items 2, 4, 5 and 6; question 21, items 2 and 5. The items were reversed that high values correspond to a high level of self-esteem. The scale value with a possible range of 1 to 4 is the mean value of the items. At least three items had to have valid values.

## 8.31 Optimism

The scale *optimismus_11* grasps the degree in optimism also called the affective valence of future perspective (Brandtstädter & Wentura 1994). It uses the following five items of the drop-off: question 2, items 2 and 5; question 16, item 1; question 21, items 1 and 6. If necessary, the items were reversed that high values correspond to a high optimism. The individual scale value with a possible range of 1 to 4 is the mean value of the items. For the scale at least three items had to have valid values.

## 8.32 Self-Efficacy

The field is covered by two generated variables.

The scale *selbstwirk_11* shows a person's degree of the self-efficacy according to Schwarzer & Jerusalem (1995, 1999). The basis are the following five drop-off items: question 2, item 7; question 16, items 3 and 7; question 21, item 3 and 4. The items were reversed that high values correspond to a high level of self-efficacy. The scale value with a possible range of 1 to 4 is the mean value of the items. At least three items had to have valid values.

The *hope_11* scale depicts the HOPE-scale of Snyder et al. (1991). The data used to form the scale are derived from several questions in the drop-off (question 2: items 2, 5 and 7, question 12: items 3 and 8, question 16: items 3, 6 and 8). The mean value of the corresponding items was calculated. A prerequisite for the formation of a mean value is the answer of at least three of the scale-relevant items. High values in the variable *hope_11* correspond to high level of self-efficacy.

## 8.33 Flexible Goal Adjustment

The scale *flex_11* measures the degree of flexible goal adjustment according to Brandtstädter & Renner (1990). It uses all ten items of question 48 in the drop-off. The items were reversed that high values correspond to a high level of goal adjustment and coping. The scale value with a possible range of 1 to 5 is the mean of the items.

## 8.34 Life satisfaction

The scale to measure life satisfaction is based on five items of question 3 of the drop-off questionnaire (Pavot & Diener 1993). At least three of the five items require valid values to calculate the mean value. High values of variable *lz_11* indicate a high level of life satisfaction.

## 8.35 Positive and negative affect

Basis for both construct of positive *(pa_11)* and negative *(na_11)* affect are the information of question 4 of the drop-off (scale by Watson, Clark & Tellegen 1988). The mean value of the positive and negative affect are constituted by the requested adjectives. At least answers to three of the items are required to calculate the index. High values indicate higher level of negative or positive affect.

## 9   ADDITIONAL VARIABLES

## 9.1   Interviewer's information of the interview situation

Following the personal interview, the interviewer gives information to the course of the interview. Some selected information is stored in the variables gci1i_1 to igci3i in the SUF DEAS 2011. The documentation of the questions is to be found directly after the CAPI presentation in the instruments. Information not provided in the SUF is available in accordance with the support of the FDZ-DZA.

| Variable | Label | Categories |
|---|---|---|
| gci1i_1 | Present at interview: no other person | Not mentioned; mentioned |
| gci1i_2 | Present at interview: spouse/partner | Not mentioned; mentioned |
| gci1i_3 | Present at interview: children | Not mentioned; mentioned |
| gci1i_4 | Present at interview: other family members | Not mentioned; mentioned |
| gci1i_5 | Present at Interview: other person | Not mentioned; mentioned |
| gci2i | Interview was translated by third person | Yes, nearly completely; yes, sometimes; no; interviewee is native speaker |
| gci3i | A present person intervened into the interview | Yes, sometimes; yes, often; no |

## 9.2 Interviewer's information on the residential environment

The interviewer captures some additional information to the residential environment besides the personal interview. A selection is stored in variables ga1 to ga4, ga7, ga12 to ga14 in the SUF DEAS 2011. They can enrich methodological analysis and enrich research questions as additional information. In detail, these variables provide the following information:

| Variable | Label | Categories |
| --- | --- | --- |
| ga1 | Location, 2011 | Single or scattered site; attached estate; peripherical location; close to centre; central location; don't know; no answer |
| ga2 | Big city with at least 100.000 inhabitants | Yes, no, not assignable |
| ga3 | Distance to centre of the nearest big town or city | *Range: 0.3 to 350 km* |
| ga4 | Prevailing building style in the street | Agricultural building; 1-2 family houses, multifamily houses, high-rise building, industrial or commercial area, not assignable |
| ga7 | Type of accommodation | One family house, multifamily home, retirement home, institutionalized elderly home, no answer |
| ga12 | Location of the interview | Respondents house or room; not respondents house or room; no interview |
| ga13 | Accessibility of house or room | Without stairs, up to 10 steps in staircase, over 10 steps in staircase, don't know |
| ga14 | Elevator on site | Yes; no; don't know |

## 9.3 Documenting social relations of respondents

### 9.3.1 Personal codes

The social relations of the interviewees to other persons are listed with the help of person codes and a corresponding person card. Person codes are three-digit codes and different for

each group of persons. Normally, these are individual codes (for children, partner, grandchildren, friends), but there are also collective codes (e.g. for neighbours).

Example: If the interviewee has children, the interviewer transfers code number 201 (that is the code for the first child) together with the name of the child to the person card. Further children get the codes number 202, 203 etc. and are noted together with their names on the respective person card. With the help of these person cards (that are filled during the interview), the interviewee and the interviewer can agree to which person information are currently collected at any time during the interview.

On the basis of these person codes further attributes of the various social relations (e.g. the second child) can be clearly assigned.

### 9.3.2  Grandchildren

In the third wave of the DEAS a second person card was introduced. It contains a separate code for each grandchild. In previous waves, only the collective code 301 was noted for all grandchildren. In the survey as well as in the data there are information for the *sole grandchild*, the *basic grandchild* and the *random grandchild*. This identification is derived from the procedure during the interview. When asked about their children, also information on the grandchildren of the interviewee are collected. The number of grandchildren determines the process of collection of further information of that grandchild.

If the child of the interviewee has only own child, it is called the *sole grandchild*. All information according to the filtering rules are now collected for this grandchild.

If the child of the interviewee has more than one own child, some basic information (year of birth, sex, place of residence) for all *basic grandchildren* are collected. The CAPI-program then randomly chooses one *random grandchild* out of the pool of all children and collects further information only on this grandchild. It serves as representative for his/her siblings. A random grandchild is always a basic grandchild but never a sole grandchild.

This procedure allows a broad but also time saving collection of information about grandchildren and an unbiased selection of that grandchild with more and detailed information. Random grandchildren are only sampled for the first four children of an interviewee. From the fifth child on, only year of birth and sex of the grandchildren are collected.

## 10  OTHER

## 10.1 Codebook DEAS2011

The Codebook of the SUF DEAS 2011 containing all specifications, labels and frequencies of all variables as well as further other documentation is available on the website of the FDZ-DZA.

## 10.2 Regional data

Information and characteristics on the level of countries and independent cities can be merged to the data. The Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR) provides information on indicators of the regional level via its system called INKAR. An overview of all available characteristics is provided by the list of indicators of the BBSR.

Registered users can get some selected indicators as data file that can be merged to the survey data after consulting the FDZ-DZA. Full indicators can only be merged to the data at a specifically protected computer workplace at the DZA. Please contact the FDZ-DZA for more information.


## 10.3 Data on non-participants

There are information about 600 target persons who were not able or did not want to take part in the questionnaire of the DEAS wave 2011. Some of this information are directly collected from the target person, other information are proxy information given by other persons. The data comprises information on the living conditions, activity status, health status as well as subjective evaluation of central living areas. Interested researchers of the short questionnaire can contact the FDZ-DZA (fdz@dza.de).

# LITERATURE

Blom, A. G., & Schröder, M. (2011). Sample composition 4 years on: retention in Share Wave 3. In M. Schröder (Hrsg.), Retrospective data collection in the Survey of Health, Ageing and Retirement in Europe (S. 55-61). Mannheim: MEA.

Brandtstädter, J., & Renner, G. (1990). Tenacious goal pursuit and flexible goal adjustment: explication and age-related analysis of assimilative and accommodative strategies of coping. Psychology and Aging, 5, 58-67.

Brandtstädter, J. & Wentura, D. (1994). Veränderungen der Zeit- und Zukunftsperspektive im Übergang zum höheren Erwachsenenalter: Entwicklungspsychologische und differentielle Aspekte. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, XXVI(1), 2-21.

Bundesamt für Bau-, Stadt- und Raumforschung (BBSR) im Bundesamt für Bauwesen und Raumordnung (2012). Raumabgrenzungen und Raumtypen des BBSR. Bonn: BBSR

Bullinger, M., & Kirchberger, I. (1998). Der SF-36-Fragebogen zum Gesundheitszustand. Handanweisung [SF-36-Questionnaire on health status]. Göttingen: Hogrefe.

De Jong Gierveld, J., & Van Tilburg, T. (2006). A 6-item scale for overall, emotional, and social loneliness. Confirmatory tests on survey data. Research on Ageing, *28*(5), 582-598.

Engstler, H., & Hameister, N. (2016). Deutscher Alterssurvey (DEAS): Kurzbeschreibung des Datensatzes SUF DEAS2011, Version 2.0. Berlin: DZA.

Engstler, H., & Motel-Klingebiel, A. (2010). Datengrundlagen und Methoden des Deutschen Alterssurveys (DEAS). In A. Motel-Klingebiel, S. Wurm & C. Tesch-Römer (Hrsg.), Altern im Wandel. Befunde des Deutschen Alterssurveys *(DEAS)* (S. 34-60). Stuttgart: Verlag W. Kohlhammer.

Engstler, H., & Schmiade, N. (2013). The German Aging Survey (DEAS) - A longitudinal and time-series study of people in the second half of life. Schmollers Jahrbuch, 133(1), 97-107. doi: https://doi.org/10.3790/schm.133.1.97

Engstler, H., & Wurm, S. (2006). Datengrundlagen und Methodik. In C. Tesch-Römer, H. Engstler & S. Wurm (Hrsg.), Altwerden in Deutschland. Sozialer Wandel und individuelle Entwicklung in der zweiten Lebenshälfte (S. 47-83). Wiesbaden: VS Verlag für Sozialwissenschaften.

Erikson, R., Goldthorpe, J. H., & Portocarero, L. (1979). Intergenerational Class Mobility in three Western European Societies. British Journal of Sociology, 30, 415-441.

Erikson, R., & Goldthorpe, J. H. (1992). The constant flux: a study of class mobility in industrial societies. Oxford: Clarendon.

European Central Bank (2016). The Household Finance and Consumption Survey (HFCS): Methodological report for the second wave, (ECB Statistical Paper Series No. 17). Frankfurt a. M.

Ganzeboom, H. B. G., De Graaf, P. M., & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. Social Science Research, 21(1), 1-56.

Ganzeboom, H. B. G., & Treiman, D. J. (1996). Internationally comparable measures of occupational status for the 1988 International Standard Classification of Occupations. Social Science Research, 25(3), 201-239.

Ganzeboom, H. B. G., & Treiman, D. J. (2003). Three internationally standardised measures for comparative research on occupational status. In J. H. P. Hoffmeyer-Zlotnik & C. Wolf (Eds.), Advances in cross-national comparison: A European working book for demographic and socio-economic variables (pp. 159-193). New York u.a.: Kluwer Academic.

García Villar, Jaume & Quintana-Domeque, Climent, (2009). Income and body mass index in Europe. Economics & Human Biology, 7(1), pp. 73-83.

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). Bamberg, Germany: Leibnitz Institute for Educational Trajectories, National Educational Panel Study.

Hautzinger, M., & Bailer, M. (1993). ADS - Allgemeine Depressions-Skala. Weinheim: Beltz.

Hoffmeyer-Zlotnik, J. H. (2003). "Stellung im Beruf" als Ersatz für eine Berufsklassifikation zur Ermitlung von sozialem Prestige. ZUMA-Nachrichten, 27(53), 114-127.

Hoyer, W. J., Stawski, R. S., Wasylyshyn, C., & Verhaeghen, P. (2004). Adult age and digit symbol substitution performance: A meta-analysis. Psychology and Aging, 19(1), 211-214.

infas Institut für angewandte Sozialwissenschaft GmbH. (2012). Methodenbericht Deutscher Alterssurvey (DEAS): 4. Befragungswelle - Panelbefragung 2011. Bonn: infas.

infas Institut für angewandte Sozialwissenschaft GmbH. (2018). Methodenbericht Deutscher Alterssurvey (DEAS): Durchführung der 6. Erhebungswelle 2017. Bonn: infas.

Klaus, D., & Engstler, H. (2017). Daten und Methoden des Deutschen Alterssurveys. In K. Mahne, J. K. Wolff, J. Simonson & C. Tesch-Römer (Eds.), Altern im Wandel: Zwei Jahrzehnte Deutscher Alterssurvey (pp. 29-45). Wiesbaden: Springer VS. doi: https://doi.org/10.1007/978-3-658-12502-8_2

Klaus, D., Engstler, H., Mahne, K., Wolff, J. K., Simonson, J., Wurm, S., & Tesch-Römer, C. (2017). Cohort profile: The German Ageing Survey (DEAS). International Journal of Epidemiology, 46 (4), pp.1-8. doi: https://doi.org/10.1093/ije/dyw326

Kohli, M., Künemund, H., Motel, A., & Szydlik, M. (2000). Soziale Ungleichheit. In M. Kohli & H. Künemund (Hrsg.), Die zweite Lebenshälfte - Gesellschaftliche Lage und Partizipation im Spiegel des Alters-Survey (S. 318-336). Opladen: Leske + Budrich.

Mahne, K., Wolff, J. K., Simonson, J., & Tesch-Römer, C. (Eds.). (2017). Altern im Wandel. Zwei Jahrzehnte Deutscher Alterssurvey (DEAS). Wiesbaden: Springer VS. doi: https://doi.org/10.1007/978-3-658-12502-8

Mayer, K.-U., & Wagner, M. (1999). Lebenslagen und soziale Ungleichheit im hohen Alter. In K.-U. Mayer & P. B. Baltes (Hrsg.), Die Berliner Altersstudie (S. 251-275). Berlin: Akademie Verlag.

Motel-Klingebiel, A., Wurm, S., Engstler, H., Huxhold, O., Jürgens, O., Mahne, K., et al. (2009). Deutscher Alterssurvey: Die zweite Lebenshälfte. Erhebungsdesign und Instrumente der dritten Befragungswelle (DZA Diskussionspapier Nr. 48). Berlin: Deutsches Zentrum für Altersfragen.

Motel-Klingebiel, A., Wurm, S., & Tesch-Römer, C. (Hrsg.) (2010). Altern im Wandel. Befunde des Deutschen Alterssurveys (DEAS). Stuttgart: Kohlhammer.

Nunn, A. J., & Gregg, I. (1989). New regression equations for predicting peak expiratory flow in adults. British Medical Journal, 298, 1068-1070.

Pavot, W., & Diener, E. (1993). Review of the Satisfaction With Life Scale. Psychological Assessment, 5(2), 164-172.

Rose, D., & Harrison, E. (Eds.) (2010). Social class in Europe. An introduction to the European Socio-economic Classification. London/New York: Routledge.

Rose, D., Harrison, E., & Pevalin, D. (2010). The European Socio-economic Classification: a prolegomenon. In D. Rose & E. Harrison (Eds.), Social class in Europe. An introduction to the European socio-economic classification (pp. 3-38). London/New York: Routledge.

Rosenberg, M. (1965). Society and the adolescent self-image. Princeton, NJ: Princeton University Press.

Schwarzer, R., & Jerusalem, M. (1995). Generalized Self-Efficacy Scale. In J. Weinman, S. Wright, & M. Johnston (Eds.), Measures in health psychology: A user's portfolio. Causal and control beliefs (pp. 35-37). Windsor, UK: nferNELSON.

Schwarzer, R., & Jerusalem, M. (1999). Skalen zur Erfassung von Lehrer- und Schülermerkmalen. Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen. Berlin: Freie Universität.

Snyder, C. R., Harris, C., Anderson, J., Holleran, S. A., Irving, L. M., & Sigmon, S. T. (1991). The will and the ways: Development and validation of an individual-differences measure of hope. Journal of Personality and Social Psychology, 60, 570-585.

Tesch-Römer, C., Wurm, S., Hoff, A., & Engstler, H. (2002). Die zweite Welle des Alterssurveys. Erhebungsdesign und Instrumente *(*DZA Diskussionspapier Nr. 35). Berlin: Deutsches Zentrum für Altersfragen.

Tesch-Römer, C., Engstler, H., & Wurm, S. (Hrsg.) (2006). Altwerden in Deutschland. Sozialer Wandel und individuelle Entwicklung in der zweiten Lebenshälfte. Wiesbaden: VS Verlag für Sozialwissenschaften.

Tewes, U. (1994). HAWIE-R: Hamburg-Wechsler-Intelligenztest für Erwachsene, Revision 1991. Handbuch und Testanweisung. Bern u.a.: Huber.

Trappmann, M. (2014). Weights. In: Bethmann, A., Fuchs, B. & Wurdack, A. (Eds.) (2013): User guide „Panel Labour Market and Social Security" (PASS). Wave 6. FDZ Datenreport 07/2013 (S. 81-99). Nürnberg.

Treiman, D. J. (1977). Occupational prestige in comparative perspective. New York: Academic Press.

Vogel, C., Wettstein, M., & Tesch-Römer, C. (Eds.). (2019). Frauen und Männer in der zweiten Lebenshälfte. Älterwerden im sozialen Wandel. Wiesbaden: Springer VS.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. Journal of Personality and Social Psychology, 54, 1063-1070.

Wechsler, D. (1955). Manual for the Wechsler Adult Intelligence Scale. New York: Psychological Corporation.

Wegener, B. (1988). Kritik des Prestiges. Opladen: Westdeutscher Verlag.

Wirth, H., & Fischer, A. (2008). ESeC - European Socioeconomic Classification. Die Operationalisierung von ESeC im kumulierten ALLBUS 1980-2006. Mannheim: GESIS. Open access: https://nbn-resolving.org/urn:nbn:de:0168-ssoar-206921