DZA | Deutsches Zentrum für Altersfragen
Forschungsdatenzentrum

April 2025

What's new? An Introduction to the 2021 Refreshment Survey of the German Ageing Survey (DEAS),
DEAS Version DOI 10.5156/DEAS.1996-2023.M.001

Version 1.0

Stefan Stuth & Nicole Hameister

Citation suggestion:

Stefan Stuth & Nicole Hameister (2025): What's new? An introduction to the 2021 Refreshment Survey of the German Ageing Survey (DEAS) (DEAS version: DOI 10.5156/DEAS.1996-2023.M.001). Berlin: German Centre of Gerontology.

German Centre of Gerontology (DZA)
Research Data Centre (FDZ-DZA)
Manfred-von-Richthofen-Straße 2
12101 Berlin, Germany

www.fdz-dza.de
fdz@dza.de

TABLE OF CONTENTS

1. THE GERMAN AGEING SURVEY (DEAS) - A STUDY ON THE SECOND HALF OF LIFE

The German Ageing Survey ("Deutscher Alterssurvey", DEAS) is a long-term study conducted by the German Centre of Gerontology (DZA) and funded by the Federal Ministry for Family Affairs, Senior Citizens, Women and Youth (BMFSFJ). It focusses on the changes in the life situations and ageing processes of people in the second half of life. Due to its cohort-sequential longitudinal design, its range of topics, its national representativeness with a comparatively high number of cases and the long observation period, the DEAS is a unique data source in Germany for interdisciplinary research into the life situations and life courses as well as the subjective views and life plans of ageing and old people.

Microdata of the German Ageing Survey (DEAS) are available for scientific research for non-commercial purposes only. The Research Data Centre provides researchers with access to the information collected as part of the surveys and advises them on its use. The anonymised DEAS datasets of all completed survey waves and the documentation material are available free of charge from the FDZ website (www.fdz-dza.de). For data protection reasons, a usage agreement must be concluded to get access to the data.

1.1 DEAS Refreshment Survey 2021

The baseline sample originally drawn for 2020 could not be surveyed as planned due to the Covid-19 pandemic. Instead, two panel surveys were conducted: A short written survey in summer 2020 on the consequences of the complete lockdown in Germany and a regular panel survey in winter 2020/21. The sample of address data that had already been collected for the planned baseline survey was prepared but not used. In order to not to waste the address data, a written refreshment survey was conducted on the basis of this data in winter 2021/22. The aim of this survey was to supplement the aged DEAS sample with a young cohort, as the last baseline survey had taken place in 2014. The respondents in the panel have aged, and the youngest respondents had already been 46 years old in the 2020/21 wave.

The 2021 refreshment survey was aimed exclusively at the 1975-1980 birth cohorts in order to supplement the youngest cohorts. This initial survey has been a paper questionnaire. The refreshment survey was conducted under special circumstances:

- This is the first initial written survey in the history of the DEAS.
- The sampling was based on a new municipality (Gemeinde) sample (clustered sample design).
- At the end of the questionnaire, respondents were asked to provide their telephone number so that they could be included in the 2023 panel survey.

The survey instrument was based in part on the short PAPI survey in summer 2020 and the regular drop-off. As in the previous waves, the fieldwork, data collection and data weighting were carried out by the Institut für angewandte Sozialwissenschaft GmbH (infas), Bonn. Many thanks go to Luisa Bischoff, who was responsible for the preparation of the refresher sample data.

1.2 DEAS Sampling

Since 1996, new base samples have been drawn every six years from the 40 to 85-year-old resident population in private households as stratified and clustered random samples - by age group, gender and region (East/West) - from the population registers and interviewed in person by interviewers. From wave 2 (2002) onwards, participants from the previous baseline surveys who were willing to participate in the panel have been interviewed repeatedly Excluded from this are the respondents from the separately drawn sample of non-German respondents in 2002 who were not included in the longitudinal sample. After the 2008 survey, the interval between the panel waves was shortened from six to three years.

In wave 1 (1996), the selection was limited to persons with German citizenship. In wave 2 (2002), two separate samples of Germans and non-Germans were drawn for the baseline survey. Since 2008, the basic sample has been drawn from the population of 40- to 85-year-olds in private households (place of main residence) without regard of nationality.

The refreshment survey was a baseline survey and included 4,092 individuals born between 1975 and 1980. Names and contact details were taken from the original residents' registration office sample intended for the cancelled 2020 baseline survey. In total, valid and completed questionnaires are available for 1,052 individuals aged between 40 and 47.

Approximately one out of eight questionnaires could not be delivered due to address issues (12.4 per cent). This was not unexpected, as the addresses had already been provided by the registration offices in autumn/winter 2019/2020. Given that up to two years had passed between the address sampling and the dispatch of survey materials, some of the intended respondents had moved in the meantime.

The response rate was 25.7 percent, while 71.8 percent of the 1,052 participants were willing to participate in the upcoming panel. Two thirds of the respondents completed the questionnaire in November and December 2021 (68.4%), with the remainder completing it in the first months of 2022.

1.3 Topics of the Refreshment Survey 2021

Due to the written survey mode, the study participants could not be asked comprehensive questions about their life situation, in contrast to regular personal DEAS interviews and the regular additional drop-off questionnaire. The questionnaire comprised 24 pages and 69 questions. Among other things, information was collected on the following topics:

- Partnership and family
- social networks and social support
- housing and mobility
- subjective well-being
- health and loneliness
- attitudes, norms and values
- Covid-19
- self-esteem, optimism and self-efficacy
- basic socio-demographic data.

1.4 Data weighting of the Refreshment Survey

The adjusted design weights of the refreshment sample were calculated in two steps: the design weight and the adjustment. The design weight consists of the inverse of the product of the selection probability of the municipalities with the selection probability of the individuals in the municipalities. The adjustment of the design weights is based on a model which calculates the probability of participation of the new sample in the 2021 wave in a logistic model. The following information was included in the calculation of the model: old and new federal states; district types, grouped; year of birth; gender.

As the 2021 refreshment sample is a baseline survey, there is only a cross-sectional weight for this dataset available (the adjusted design weight - *qs_a21*).

The use of weights is at the responsibility of the individual user. Data weighting is one of several ways of dealing with different sampling probabilities and the problem of selective participation.

2 DATA FORMATS

The Scientific Use Files (SUF) for the years 1996 to 2023, the weights dataset and the SUF Meta (see Section 6) are available in both SPSS and Stata data format. If there are any version-related problems with importing the data, please contact the FDZ team directly (fdz@dza.de). Most data sets and the corresponding documentation are available in German and English, the remaining data sets and documentation will be provided soon in English as well.

3 CLASSIFICATION OF VARIABLE NAMES & MISSING CODES IN THE REFRESHMENT SAMPLE

The 2021 refreshment sample also follows the variable naming system introduced in 2008. In the first two DEAS waves, variable names were still based on a card system coding tradition[1].

To make variables easier to identify, they have been consecutively assigned numbers according to the question number since the third wave. The question numbers are grouped thematically, for example, questions starting with 300 contain all information related to the respondents` family situation.

The variables from the 2021 refreshment survey begin with a ‘k’ as a wave identifier, followed by a ‘p’ to indicate the mode of data collection (paper-based survey)[2]. Due to the temporal proximity of the surveys and the absence of respondent overlap, both the 2020/21 panel wave and the 2021 refreshment sample use the same wave identifier ‘k’. Variables from the 2020/21 panel survey are marked with ‘kc’ or ‘kd’, whereas variables in the 2021 refreshment sample begin with ‘kp’. The number that follows corresponds to the corresponding question number.

In the case of multiple responses, the answer options are distinguished by a consecutive numerical suffix (e.g. kp34_1, kp34_2, etc.). Open-ended responses are identified by the suffix ‘o’ (e.g. kp9o).

A complete overview of all variable names ever used in the DEAS datasets is available in the variable correspondence list on the FDZ-DZA documentation page. This page also includes the

[1] For example, question 101 on employment status from the personal interview in the second wave in 2002 (labelled with ‘w’) appeared on card 14 in second position and is therefore named w14_12.
[2] Regular DEAS waves use either a ‘c’ or ‘d’ to indicate the type of survey conducted (‘c’ for computer-assisted personal interview; ‘d’ for drop-off questionnaire, either written or web-based).

codebook, which provides the basic frequencies for all variables available in the SUF DEAS refreshment sample 2021, including the corresponding labels.

The assignment of missing codes follows the revised system introduced in wave 4. This improves the understanding of the reasons behind non-valid data.

The following table describes the missing value codes used in the data.

SPSS	Stata	Label [Explanation]
-1	.a	Denied
-2	.b	Don't know
-3	.c	Filtered (question)
-4	.d	Filtered (sample) *[if the question is only aimed at panel or baseline respondents; not applicable in the 2020/21 wave]*
-5	.e	No Drop-off available
-6	.f	Not specified *(Drop-off)*
-7	.g	Deleted in data edition

In general, missing codes in the dataset are defined as user-defined system missings. Exceptions apply to waves 1 to 3 (1996, 2002, and 2008), in which this system for labelling missing values had not yet been introduced. For these waves, it is recommended to run frequency counts of the variables to be analysed prior to analysis, in order to identify missing codes and to define them accordingly as missing values.

4 MODIFICATION OF DATA IN THE SUF TO MAINTAIN DE FACTO ANONYMITY

In order to preserve the de facto anonymity of the survey participants, individual variables were not included in the Scientific Use File (SUF) of the 2021 refreshment sample; in addition, a number of variables were altered or manually changed in open-ended responses.

- Almost all variables relating to spatial context were removed. Information on respondents' month and year of birth was also deleted.
- Individual variables containing open-ended responses were either deleted or edited to generalise highly specific details (such as years or locations). These changes are marked with a '#' prefix in the value labels.
- Rare family demographic characteristics were top-coded.
- Rare annual data were grouped into broader categories ("top-coding") by recoding values within the existing numerical variable to either a central value in the new range or to the modal value. These were relabelled accordingly.
- In some cases, maximum values were combined into an open-ended top category (e.g. for household size).

The DZA has set up specially secured PC workstations for guest researchers, which allow them to carry out analyses with the complete DEAS datasets and additional information, e.g. for small-scale analysis at district or municipality level or including contextual characteristics of the immediate neighborhood (Microm data). If you are interested, please contact the FDZ team: fdz@dza.de.

5 THE COMPLEX DEAS SURVEY DESIGN

The DEAS uses a complex stratified and clustered random sampling method to ensure that the sample is representative and reduces both random and systematic biases. Here's a breakdown of the methodology:

1. Stratification: Before sampling, the population is divided into different strata based on:
    - Gender
    - Three age groups
    - Region (old or new federal states) Additionally, people from the oldest age group and the new federal states are oversampled to ensure adequate representation of these groups.
2. Clustering: To make the survey logistically feasible and cost-effective, it is clustered. Instead of randomly selecting individuals from the entire country, the sample is drawn from specific municipalities, which are divided into two groups: those from the old federal states and those from the new federal states. Within each group, municipalities are randomly selected based on size, meaning larger municipalities are more likely to be selected.
3. Sampling Process:
    - Proportional to size: Larger municipalities have a higher chance of being selected.
    - From 11,014 municipalities: A sample of 59 municipalities from the new federal states and 142 from the old federal states was drawn in 2020.
4. Why it matters:
    - Clustering and dependence: Respondents within the same community may share similar characteristics, which can lead to dependencies between observations. For instance, people in wealthy areas may have similar income levels.
    - Impact of clustering on analyses: If the clustered sampling method isn't taken into account, the analysis would treat respondents as independent, which could lead to a wrong estimation of variance and potentially produce misleading results.

The svyset command in Stata is used to define the survey design for complex survey data, taking into account stratification, clustering, and weights.

```
svyset psu, strata(strata_psu) fpc(fpc_1) vce(linearized)
singleunit(scaled) || ssu [pw=weight], strata(strata_ssu)
```

- psu: This specifies the primary sampling units (PSUs), which are the clusters of the survey sample. In the case of DEAS, this corresponds to the municipalities or other units of clustering.
- strata(strata_psu): This option defines the strata for the primary sampling units (PSUs). These strata are defined by region (old or new federal states).

- fpc(fpc_1): This refers to the finite population correction (FPC) factor, which adjusts for the fact that the primary sampling units are drawn from a finite population. Whereas the impact of the correction is usually negligible it comes with a great advantage: This option ensures that the variance is correctly adjusted when multiple waves of data are considered, where respondents repeatedly took part.
- vce(linearized): This option specifies the method for variance estimation. The linearised option is commonly used in survey data to compute robust standard errors that account for the complex design of the survey, including clustering and stratification.
- singleunit(scaled): This option adjusts for single units (clusters) that have only one observation (for example, if a municipality has only one respondent). The scaled option ensures that the variance for these units is correctly calculated.
- || ssu: This part indicates the secondary sampling units (SSUs) that are the individual respondents, which were sampled within each PSU.
- [pw=weight]: This specifies the weight variable that should be used to adjust for unequal probabilities of selection, nonresponse, and other design features. The weight variable adjusts the analysis to ensure that the sample is representative of the population. You will need to create or select the appropriate weight variable.
- strata(strata_ssu): This option defines the strata at the level of the secondary sampling units (SSU) that are defined by gender and age groups, and (implicitly) region.

The appropriate weight must be selected or created independently according to the analysis design. All subsequent analyses that include the suffix -svy take into account the complex survey design of the DEAS, which was defined using the svyset command.

If fpc_1 is also taken into account, Stata also knows that the same SSU (the respondents) occur several times (e.g. if several cross-sections are analyzed). The variance calculations are corrected accordingly for analyses with svy commands.

In summary, the svyset command is crucial for correctly handling the DEAS survey design in Stata, ensuring that the data is analysed properly by accounting for complex design features such as stratification and clustering.

6 DOCUMENTATION

In the DEAS documentation section of the FDZ-DZA website, you can view and download a range of information and materials related to all available DEAS waves. These include:

- Complete Survey Instruments: The full questionnaires used in each wave of the DEAS, providing insight into the exact questions asked during data collection.
- Method Reports (infas): Detailed reports on the methodology used in the survey, explaining the sampling methods, data collection procedures, and other methodological considerations.
- Codebooks: A comprehensive listing of all the variables collected in each wave, along with their coding and definitions, essential for understanding how the data is structured.
- Working Aids: Tools to assist with specific aspects of the data, such as guides on interpreting characteristics like grandchild data.
- Variable Correspondence List (Excel Format): This document helps users identify how variables correspond across different waves, crucial for longitudinal analysis.
- Detailed Variable Documentation: In-depth descriptions of each variable, including its definition, coding, and any important considerations when analysing it.
- Explanation of Data Weighting: Information on how to apply the appropriate weights to ensure the data is representative and your analyses are methodologically sound.
- Publications based on DEAS: The website also offers a catalogue of publications based on DEAS data, available in a complete PDF format or as an interactive list with filter options, making it easier to find relevant research.

Most of these documents are available in English, allowing broader access for international researchers.

These resources are invaluable for anyone working with DEAS data, offering clear guidance on how to use the data correctly and ensuring your analyses are both accurate and rigorous.

LITERATURE

Federal Office for Research on Building, Urban Affairs and Spatial Development (BBSR) in the Federal Office for Building and Regional Planning (2012). Spatial delineations and spatial types of the BBSR. Bonn: BBSR. URL: (https://www.inkar.de/) [last viewed on 01.05.22].

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.: European Central Bank. URL: https://www.ecb.europa.eu/pub/research/statistics-papers/html/index.en.html [last viewed on 01.06.22].

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: Leibniz Institute for Educational Trajectories, National Educational Panel Study. URL: https://www.neps-data.de/Datenzentrum/Publikationen/NEPS-Survey-Papers [last viewed on 01.06.22].

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), 1105-1105g. doi: https://doi.org/10.1093/ije/dyw326

Mahne, K., Wolff, J. K., Simonson, J., & Tesch-Römer, C. (Eds.) (2017). Ageing in transition. Two decades of the German Ageing Survey (DEAS). Wiesbaden: Springer VS. Open Access: doi: https://doi.org/10.1007/978-3-658-12502-8.

Schiel, S., Ruiz Marcos, J., Schulz, S., Dickmann, C. & Fahrenholz, M. (2021). Methodological report. German Ageing Survey (DEAS): Implementation of the 7th survey wave 2020/21. Bonn: infas. URL: https://www.dza.de/forschung/fdz/fdz-publikationen [last viewed 01.06.22].

Trapmann, M. (2014). In: Bethmann, A., Fuchs, B., Wurdack, A., Beste, J., Trappmann, M., Müller, G., Eggs, J., Gundert, S., Wenzig, C., Gebhardt, D. (eds.) (2013). User guide "Panel Study Labor Market and Social Security" (PASS) * wave 6. (FDZ-Datenreport, 07/2013 (en)), Nuremberg: IAB (pp. 81-99). URL: https://fdz.iab.de/187/section.aspx/Publikation/k131218302 [last viewed on 01.06.22].

Vogel, C., Wettstein, M., & Tesch-Römer, C. (Eds.) (2019). Women and men in the second half of life. Ageing in social change. Wiesbaden: Springer VS. Open Access doi: https://doi.org/10.1007/978-3-658-25079-9