Data and Methods Report
October 2026

Andreas Woisch | Barbara Franke | Heiko Quast | Dennis Föste-Eggers | Hanna Mentges | Ann-Christin Renneberg | Laura Schoger | Anna Meine | David Nika | Daniel Klein | Anne Weber | Thorsten Euler | Henrike Steins

DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018

Data and Methodology Report on the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018
(1st–3rd survey waves)

fdz.DZHW.
Forschungsdatenzentrum
für Hochschul- und Wissenschaftsforschung

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International Licence (CC-BY-NC-SA)

https://creativecommons.org/licenses/by-nc-sa/4.0/deed.en

Project Lead

Dr Johann Carstensen
Telephone +49 511 450670-914
Email: carstensen@dzhw.eu

Dr Heike Spangenberg
Telephone +49 511 450670-151
Email: spangenberg@dzhw.eu

Project team members
Heiko Quast, Sebastian Lang, David Nika, Barbara Franke, Verena Jahn, Hanna Mentges, Laura Schoger, Dennis Föste-Eggers

Legal notice

Published by
German Centre for Higher Education Research and
Science Studies GmbH (DZHW)
Lange Laube 12 | 30159 Hanover | www.dzhw.eu
PO Box 2920 | 30029 Hanover
Tel.: +49 511 450670-0 | Fax: +49 511 450670-960

Management
Dr Marcus Beiner
Regina Oelfke

Chairman of the Supervisory Board
Head of Section Peter Greisler

Register court
Hanover Local Court | B 210251
VAT registration number:
DE291239300

October 2026

Table of Contents

List of Tables 3

List of Figures 3

I Introduction 4

1 Content and Structure of the Study 5

2 Survey Instrument 8

2.1 Contents of the Survey Instruments 8

2.1.1 Wave 1 8

2.1.2 Wave 2 9

2.1.3 Waves 3a and 3b 9

2.2 SLC Episode Module 10

2.3 Components of the Survey Instruments used in a Secondary Capacity 11

2.4 Pre-Test 11

2.4.1 Wave 1 12

2.4.2 Wave 2 12

3 Population, Sample Size and Response Rate 14

3.1 Wave 1 15

3.2 Wave 2 19

3.3 Wave 3a 21

3.4 Wave 3b 22

4 Data Preparation 24

4.1 Data Transfer 24

4.2 Coding of (semi-)open Responses 24

4.3 Data Checking and Cleaning 25

4.4 Generation of Variables 27

4.5 Coding of Missing Values 27

4.6 Generation of the Dataset 28

4.7 Assignment of Variable Names, Variable Labels and Value Labels (across Waves) 29

5 Weighting 30

5.1 Procedure and Application Notes 30

5.2 Weighting the Dataset 31

5.2.1 Wave 1 31

5.2.2 Waves 2, 3a and 3b 32
6 Anonymisation 34
7 References 36
Appendix 37

List of Tables

Table 1: Gross and net sample sizes and response rates for the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018 18
Table 2: Coded variables and code lists in the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018 25
Table 3: System of missing values in the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018 27
Table 4: Weights provided for the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018 30
Table 5: Modes of access at the Research Data Centre for Higher Education Research and Science Studies (FDZ-DZHW) 34
Table 6: Documentation of the origin of reused questions in Wave 1 (student survey) 37
Table 7: Documentation of the origin of reused questions in Wave 1 (headteacher survey) 39
Table 8: Documentation of the origin of reused questions in Wave 2 40
Table 9: Documentation of the origin of reused questions in Waves 3a and 3b 42

List of Figures

Figure 1: Cohort-panel design of the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification Series 6
Figure 2: Survey schedule for DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018 6
Figure 3: Response rate of the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018 over time, Wave 1 18
Figure 4: Response rate of the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018 over time, Wave 2 21
Figure 5: Response rate of the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018 over time, Wave 3a 22
Figure 6: Response rate of the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018 over time, Wave 3b 23

I Introduction

The DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification are a series of studies examining the educational, career and life paths of school leavers with higher education entrance qualifications. They are conducted by the German Centre for Higher Education Research and Science Studies GmbH (DZHW), funded by the Federal Ministry of Education and Research (BMBF), and – complementing official higher education statistics – serve the purposes of national education monitoring and addressing research questions in the fields of higher education and science studies. Selected cohorts of school leavers have been surveyed since 1976.

The data from some cohorts of school leavers with higher education entrance qualifications are processed and documented for secondary use. They are made available via various modes of access as Scientific Use Files (SUF) for secondary scientific use. In addition to the survey datasets, documentation relating to the datasets and the survey administration are also provided. This data and methods report forms part of the documentation for the first to third survey waves of the 2018 DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification[1]. Further documentation relating to the study (data set reports, questionnaires, question flow diagrams, etc.) can be downloaded free of charge from the FDZ-DZHW search portal[2].

Chapter 1 provides a general overview of the content and structure of the 2018 DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification. The rest of the report is structured primarily according to the sequence of the research process. Chapter 2 describes the survey instruments used, whilst Chapters 3 and 4 describe the survey process (survey procedure, response rate, data preparation). Chapters 5 and 6 then outline the weighting and anonymisation procedures carried out.

[1] This document is the English translation of the original German report. There may be instances where certain terms or phrases could not be translated (including in some tables or figures). We apologise for any inconvenience this may cause.
[2] https://metadata.fdz.dzhw.eu/en/start

1 Content and Structure of the Study

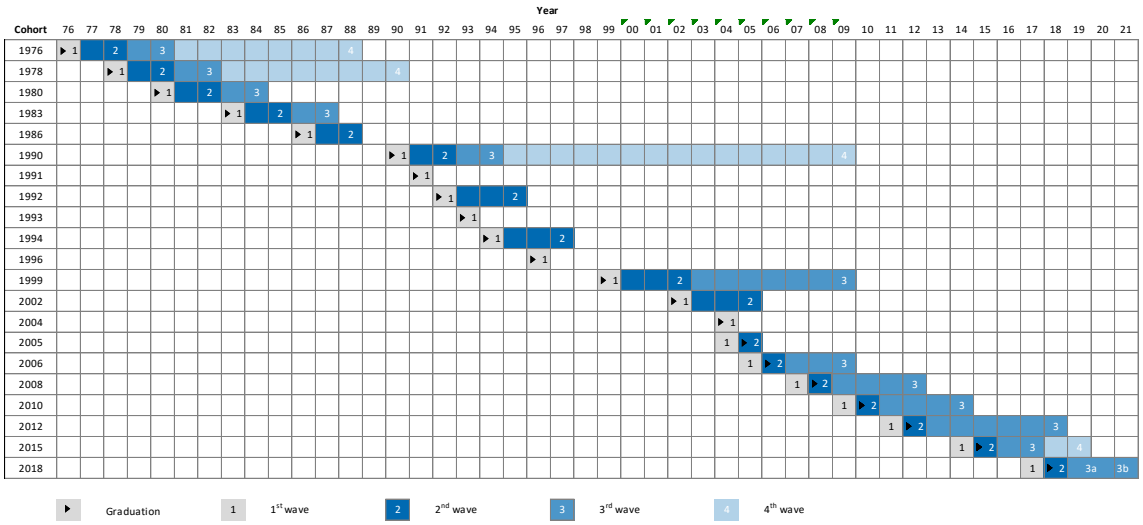
[Survey series] The 2018 DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification forms part of the DZHW series of surveys of higher education-eligible students, in which standardised multiple-wave surveys are used to gather information on the respondents' transition from school to higher education and vocational education and training, as well as their post-school educational and employment trajectories in their entirety. The first DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification was conducted in 1976; since then, 21 cohorts of students eligible for higher education have been studied.

The population of each cohort comprises school leavers with a higher education entrance qualification who, in the respective school year, attend general education schools and vocational schools in Germany where they obtain a general higher education entrance qualification (Abitur), an advanced technical college certificate (Fachhochschulreife), or a subject-related higher education entrance qualification (subject-specific higher education entrance qualification or state-specific education entrance qualification).

As a rule, several survey waves are conducted for each cohort of students eligible for higher education at different points in time, both before and after they have obtained their higher education entrance qualification. This therefore constitutes a combined cohort-panel design. Surveys of cohorts of students eligible for higher education traditionally comprised between one and four waves. Until 1986, the DZHW aimed to survey each cohort of students eligible for higher education shortly after, and 2½, 4½ and 12½ years after leaving school. From 1990 onwards, the survey on delayed or revised educational and study decisions, which was conducted two and a half years after leaving school, was discontinued. Since 2005, an additional survey has been carried out half a year before students obtain their higher education entrance qualification. This first wave of surveys is followed by a second wave for each cohort, which takes place half a year after obtaining higher education entrance qualifications. The third wave of surveys follows approximately two and a half years after graduation. Over a period of 45 years up to the present survey, this has resulted in a time series comprising 21 cohorts with a total of 55 survey waves (see

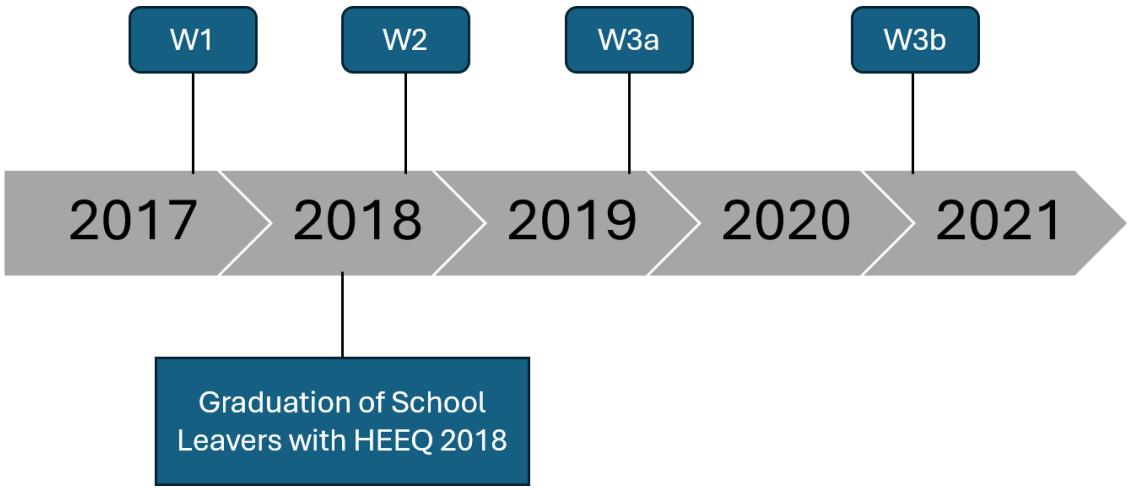
Figure1).

Figure1: Cohort-panel design of the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification Series



Unlike previous cohorts of higher education-entry-eligible students, the third wave of the 2018 cohort was surveyed in two sub-surveys: Wave 3a (interim wave) and Wave 3b (regular third wave). The reason for this was the implementation study of the Student Life Cycle Panel (SLC), within the framework of which data on higher education-entry-eligible students is to be collected in the future. One of the aims of the implementation study was to test the impact of a shorter survey interval on panel stability. Consequently, as part of a methodological experiment involving a random subsample, Wave 3a was conducted approximately 1½ years after leaving school, whilst Wave 3b was carried out 2½ years after obtaining the higher education entrance qualification. As part of the Student Life Cycle Panel, the plan is to subsequently survey school leavers with a higher education entrance qualification in 2018 at two-year intervals up to Wave 8.

Figure2: Survey schedule for DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018



[Analytical potential] To enable a comparison over time between different cohorts, an identical core set of information is surveyed, in some cases within the same survey waves or at the same time points. This allows long-term trends in post-school pathways to be examined, including through time-series and cohort comparisons. Furthermore, within a cohort, some of the questions are repeated across the various survey waves. This enables the analysis of intra-individual changes between waves (e.g. causal panel analyses). In addition, depending on current developments and research interests, certain aspects are explored in greater depth or supplemented in individual cohorts.

[Context within the field of research] The sample and survey design, as well as the associated analytical possibilities, distinguish the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification from other surveys of school leavers with a higher education entrance qualification conducted in Germany. No other survey enables nationwide analyses. Furthermore, other surveys in this field have no time series, or significantly shorter ones. Other surveys of higher education-entrance-eligible students include, for example, the Saxon Graduate Study[3] (conducted by the Competence Center for (Higher) Education Research at TU Dresden), the TOSCA study[4] (now conducted by the Hector Research Institute of Education Sciences and Psychology at the University of Tübingen) and the Berliner-Studienberechtigten-Panel[5] (Best Up, conducted by the German Institute for Economic Research (DIW) and the Berlin Social Science Centre (WZB)).

[Specific features of the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018] In addition to the general characteristics of the survey series, the 2018 cohort of school leavers with a higher education entrance qualification examined here exhibits the following specific features: Thematically, the emphasis is on diversity, with a particular focus on migration background, as well as on the subjective health and well-being of the school leavers with a higher education entrance qualification. Furthermore, for the first time within this survey series, school heads are being questioned on various topics, such as the composition of the student body, how schools deal with diversity, spatial, human and financial resources, and measures relating to career and higher education guidance.

As part of the second wave, a survey experiment was also implemented to analyse the effects of the survey method on participation rates in a panel survey using either a written postal questionnaire or an online survey (see 3.2).

3 See also: https://tu-dresden.de/zqa/forschung/forschungsprojekte/absolventenstudien/saechsische-absolventenstudie/index?set_language=en (accessed on 30 October 2026)
4 See also: https://uni-tuebingen.de/en/faculties/faculty-of-economics-and-social-sciences/subjects/department-of-social-sciences/education-sciences-and-psychology/research/current-studies/tosca/ (accessed on 30 October 2026)
5 See also: https://www.wzb.eu/en/research/dynamics-of-social-inequalities/skill-formation-and-labor-markets/best-up (accessed on 30 October 2026)

2 Survey Instrument

The first survey wave of the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018 was conducted as a standardised self-administered survey (Paper and Pencil Interview; PAPI), in which a standardised paper questionnaire was used as the survey instrument for both students and school heads. In the second survey wave, depending on the group allocation (see Chapter 3.2), either an online survey (computer-assisted web interview; CAWI) using a standardised online questionnaire or a PAPI survey using a standardised paper questionnaire was conducted. The online questionnaire was programmed and the online survey was conducted using the DZHW online survey software 'Zofar'. For both surveys in the third wave, only a standardised online questionnaire was used.

Chapter 2.1 presents the key content of the survey instruments. Chapter **Fehler! Verweisquelle konnte nicht gefunden werden.** refers to the components of the survey instruments used for secondary purposes, which are listed in the appendix. Chapter 2.3 describes the pre-tests carried out to review and improve the questionnaires. Within the chapters, the first, second and both third survey waves are each described in sub-sections.

2.1 Contents of the Survey Instruments

2.1.1 Wave 1

The questions in the first wave of the student questionnaire relate to five broad thematic areas: questions about school life, life after school, decision-making, personality and values, and personal details. The first thematic area includes questions about the type of school attended, the type of desired qualification, the average mark, and satisfaction with one's own performance and school lessons in general. There are also questions about personal strengths and weaknesses, exam anxiety and career interests. Career and educational intentions form part of the second thematic area, which includes questions on the training occupation or degree course the respondent is aiming for, their chosen higher education institution, and their parents' wishes. Decision-making is explored in the third section through questions on the respondent's current level of information regarding training and study opportunities, difficulties in choosing a post-school pathway, support received in the decision-making process, and career prospects for vocational training and degree courses. The fourth thematic area focuses on career and life goals, risk-taking, religiosity and political participation, whilst the final section includes questions on the respondents' gender, age and country of birth, as well as their book ownership, their parents' country of birth and their parents' educational qualifications.[6] It is worth noting that the first survey wave specifically incorporated survey instruments designed to capture the processes and mechanisms that explain the actual decisions, educational trajectories and educational outcomes (e.g. social and political engagement) recorded in subsequent waves. The operationalisation of constructs from various theoretical approaches aimed at explaining educational and career decisions was therefore a key focus of the first survey wave. As the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018 also regards the issue of diversity, with a particular focus on migration background, as a key priority, surveys were conducted on both indicators for determining migration background and attitudes towards the situation of migrants in Germany.

[6] Due to data protection regulations, no information on the parents of respondents from the states of Hesse and Berlin could be collected in the first wave (see chapter 3). However, the relevant questions were asked in the second wave for respondents from Hesse and Berlin.

For the first time in the survey series, the first wave also included a supplementary survey of school heads using a separate questionnaire on various topics, such as the composition of the student body, how schools deal with diversity, spatial, human and financial resources, and measures relating to career and higher education guidance.

2.1.2 Wave 2

In the second wave, the questions can be categorised into a total of seven broad thematic areas: questions on school years, the application process, post-school pathways, choice of degree course and career choice, living circumstances, leisure and health, as well as personal and family details. The first thematic area includes questions on the type of university entrance qualification, the date on which it was obtained, the final grade, the most recent school report grades in various subjects, and any previous vocational education and training or employment. In addition, respondents are asked about the examination subjects (including the number of lessons and the teacher's gender) and their interest in those subjects. For the second section of the application process, regarding vocational education and training, data is collected on the number of different training organisations and training applications, as well as the names of the training occupations. Where respondents had applied for a degree programme, corresponding questions were asked regarding the number of higher education institutions, subjects or subject combinations, and the names of the degree programmes, subjects and higher education institutions. The decision regarding vocational training or higher education forms part of the third thematic area, which covers the (main) occupation in December 2018, the next planned steps, the higher education institution, the degree programme and subjects studied, the type of degree, intentions to pursue a Master's degree, the name of the vocational education and training programme, the location of the vocational school and the training company, the job title, the company size, the economic sector and the place of work – where applicable – for both current and planned employment. In addition to these core instruments, there are also questions regarding the assessment of career prospects, labour market considerations, teaching and learning environments, the assessment of one's own performance, the role of admission restrictions, stays abroad, the motives for the chosen post-secondary pathway, and the reasons why post-secondary plans may not yet have been realised. The fourth thematic area focuses on the respondent's dream job and their intended career, statements regarding knowledge and science, and an estimate of monthly income upon entering the workforce and after ten years in the profession. Regarding living circumstances, questions are asked about monthly income, moving out of the parental home, and relationships. The 'Leisure and Health' section comprises questions on leisure activities, general, mental and physical health, and stress. Finally, the last section asks about career and life goals, as well as general life satisfaction. In addition, the sociodemographic characteristics from the first wave are supplemented by language skills, the language spoken at home, parents' occupations, occupational groups and highest school leaving qualification[7], as well as grandparents' academic degrees and country of birth.

2.1.3 Waves 3a and 3b

The content of the two surveys in the third wave is described in more detail below. As an interim survey comprising six modules (including the episode module (see chapter 2.2)), Wave 3a is a significantly shorter survey than the regular survey in Wave 3b. The thematic sections in Wave 3a include questions on career prospects, the episode module, questions on vocational training and higher education, health, and sociodemographic questions. The first module contains questions on satisfaction with educational and career paths, career aspirations and prospects, and time preferences. The episode module involves the episodic recording of post-school educational and employment trajectories. Respondents who have not yet commenced vocational training are asked about their intentions regarding vocational training in the 'Vocational Training' module. The 'Higher Education' module

[7] For respondents from Hesse and Berlin, the questions about their parents – which could not be included in the questionnaire in the first wave for data protection reasons – were also asked; see footnote 6.

includes questions on subjective prospects of success and the influence of costs on the decision to undertake higher education – aimed at those who have not yet commenced a degree programme. School leavers with a higher education entrance qualification who have already commenced a degree programme are further surveyed about the initial phase of their studies, academic and social integration, their subjective assessment of their performance, learning environments and problems encountered during their studies. Subsequently, all respondents are asked about their intention to commence or continue a (master's) degree programme. Those who indicate such an intention are then asked further questions about their intended degree programme. These cover the start date, subjects, higher education institution and type of qualification for the planned programme. The 'Health' module includes questions on health status, stress levels and days taken off due to illness. The final module, 'Sociodemographics', contains questions on financial security, housing situation, gender, and life satisfaction.

Wave 3b comprises a total of eight survey modules (including the 'Episodes' module): Introduction, Covid-19, School leaving qualification, Episodes, Vocational Training, Higher Education, Health, and Sociodemographics. In addition to the content from Wave 3a, the first module includes an item on risk-taking behaviour. The Covid-19 module asks questions about experiences during the Covid-19 pandemic. These include how individuals were personally affected, as well as any problems or restrictions relating to career and educational pathways, vocational training, higher education, employment and self-employment. Respondents are also asked about their future prospects. For those who last took part in the first wave (due to temporary dropout), the school leaving qualification obtained is also surveyed. This includes information on the type of higher education entrance qualification, the date of attainment and the final grade. As in Wave 3a, the 'Episodes' module contains questions on post-school educational and employment trajectories. Due to their previous participation, study participants from Wave 3a only need to cover a shorter reference period (one year instead of 2.5 years). In the 'Vocational Training' module, questions regarding intentions to undertake vocational training, the planned start date and the chosen training occupation are asked of those who are not currently in vocational training. Respondents who have already started vocational training are asked about their satisfaction with it. School leavers with a higher education entrance qualification who have not yet commenced a degree programme are asked, in the 'Higher Education' module, about the prospects of success and the costs of studying. Study participants who have already commenced a degree programme are asked to provide details on the concept of academic aptitude, academic and social integration, subjective assessment of their performance, digital equipment, teaching and learning environments, as well as intentions to study abroad and previous stays abroad for study purposes. In addition, they are asked questions regarding a potential Master's degree, covering their level of information, reasons for and against undertaking a Master's degree, and their estimated prospects of success. Subsequently, all respondents – as in Wave 3a – are asked about their (further) intentions regarding higher education. Study participants who have already commenced a degree programme may also provide information on their intentions to pursue a doctorate. The 'Health' module is identical to Wave 3a. The final module, 'Sociodemographics', contains questions on the topics of self-efficacy, social trust, epistemological beliefs, financial security, housing situation, gender, partner, children, parents' educational and vocational qualifications, parents' countries of birth and language spoken at the parental home, as well as general life satisfaction.

2.2 SLC Episode Module

A key feature of the two surveys conducted as part of the third wave of the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification in 2018 was the testing of elements from the implementation study of the Student Life Cycle Panel (SLC) project. This included, in particular, the recording of life course dynamics using online survey systems. Due to their complexity, life

course data in comparable studies were largely not collected online, or only in a rudimentary form.[8] In the DZHW's studies – particularly the graduate surveys – an event history calendar[9] was primarily used, which, depending on the study, collected limited information on selected activities via additional questions. This event history calendar was also used in the second wave of the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018. However, the event history calendar did not allow parallel episodes or adjacent episodes of the same type to be displayed separately. Furthermore, changes within a single activity could only be adequately represented to a limited extent.

As part of the SLC planning study, an integrated episode module was designed to enable a direct link between the event history calendar and the subsequent questions regarding the specification of the episodes. The specific follow-up questions relating to the respective episode types had already been compiled in advance as part of the SLC implementation study.

Episodes were recorded using the integrated episode module as part of the two surveys in the third wave of the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification. This module comprised eight different episode types[10], which, in various combinations, made it possible to record separate episodes (loops) or changes to key characteristics within an episode (splits). If the number of technically available loops/splits for a particular episode type had been exhausted, respondents could still provide details of the most recent or current episode of that type in a final query. In the surveys, the start and end dates of episodes could only be specified within defined reference periods. Accordingly, the resulting episodes may be left- or right-censored[11]. The raw data collected were subsequently processed by the primary research project so that each loop or split is treated as a separate episode and stored in a spell data set[12].

2.3 Components of the Survey Instruments used in a Secondary Capacity

Not all the questions and items used in the surveys were developed in-house by the primary research project or were formulated in everyday language. In some cases, scales from survey instruments used in other studies were adapted for use. The relevant references are listed in the Appendix inTable7,Table8 andTable9.

2.4 Pre-Test

[Objective and procedure] The survey instruments were tested in pre-tests prior to the main survey. Firstly, the aim was to check whether the questions and answer options already used in previous cohorts were perceived in the same way by the 2018 cohort of students eligible for higher education admission as they had been by earlier cohorts.[13] Secondly, the newly introduced survey instruments were to be tested for comprehensibility, answerability, theoretical validity, reliability and validity.

[8] In the National Educational Panel Study (NEPS), for example, this information is collected via telephone interviews, which requires greater resources than online surveys.
[9] At the DZHW, this term refers to a calendar-based representation of months and years, with the option to record various types of episodes.
[10] School, vocational training, higher education studies, internship, continuing education, employment as an employee, unemployment, self-employment, and other activities (not processed or made available by the primary research project).
[11] Respondents were able to indicate an earlier start date or a later end date using additional variables.
[12] Due to the different approaches used to record life course episodes, the data package for the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification contains various versions of spell data sets (event history calendar, episode module).
[13] The survey instruments for the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification in 2018 are essentially based on the design established for the study series and the written questionnaires that have proven effective in the previously surveyed cohorts of school leavers with a higher education entrance qualification.

Thirdly, the inclusion of the new questions entailed fundamental changes to the structure and layout of the questionnaire, as well as to the duration of the survey, which had to be evaluated.

An initial version of the survey instruments was first assessed through expert reviews (cf. Häder, 2015, pp. 406–407). In addition, selected (newly developed or adapted) instruments were tested in three cognitive pre-tests. Alongside these assessments, targeted development and qualification pre-tests were carried out[14]. These were conducted as so-called 'field pre-tests' (cf. Häder, 2015, p. 396), i.e. under conditions as similar as possible to those of the actual surveys.

2.4.1 Wave 1

[Study participants] The cognitive pre-tests were conducted as internal tests involving both research assistants and student assistants at the DZHW. By conducting a survey of these study participants, experts from the field of higher education research were also involved, enabling them to evaluate the two questionnaires accordingly.

The field pre-tests were external pre-tests. In the first wave, three schools took part in the pre-test of the student questionnaire. The three schools all had different profiles (a general education school and two vocational schools with different specialisms), meaning that 152 students from different school types could be surveyed within their class groups. This ensured that the instruments, some of which were specific to particular school tracks, could also be tested within those respective tracks. The school heads' questionnaire was tested by several external teachers. The external testers were recruited through personal contacts.

[Implementation] The external pre-test phase for the first wave of the student questionnaire took place between 8 June 2017 and 14 June 2017, approximately six months before the fieldwork began. During this phase, questionnaires accompanied by instructions were sent by mail to one school. Staff from the team conducting the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification visited the other two schools and asked the test participants to complete the standardised questionnaire and to note down any comprehension difficulties, criticisms or comments. After completing the questionnaire, respondents were asked to fill in a feedback questionnaire on various aspects of the survey. This questionnaire collected information on the time taken to complete the questionnaire, its content and length, its structure and layout, the clarity of the questions and instructions, and the comprehensiveness of the response categories.

Based on the results of the pre-tests, the wording of various questions was refined, and the order of items and response categories within individual item sets and multiple responses was revised. Furthermore, the length of the questionnaire was reduced by shortening item sets, as the time taken to complete it exceeded the originally planned 30 minutes. The basic structure, however, remained unchanged.

The pre-test for the school heads' questionnaire involved three internal experts receiving the draft questionnaire and providing written feedback (received between 17 and 18 July 20217). The same procedure was followed for the external pre-test involving teachers. As a result, the wording of various questions in the school heads' questionnaire was clarified, and items and response categories in individual item sets, as well as multiple responses, were revised.

2.4.2 Wave 2

[Study participants] For the pre-test of the second wave, 71 written questionnaires returned by mail from first-semester students attending a lecture at the University of Hamburg, as well as 29 (written

[14] Unlike development pre-tests, qualification pre-tests are generally used for minor final adjustments to the questionnaires, such as shortening or rearranging items, as well as for testing new filter procedures or changes to the layout; see Schnell, Hill and Esser (2005), p. 348.

and returned by mail) questionnaires from participants in the first wave who were not part of the survey sample (e.g. because they would only be completing the school-based part of the advanced technical college certificate (Fachhochschulreife) in the reference year; see also chapter 3 on the population) were available.

[Implementation] The external pre-test phase of the second wave took place in November 2018. The questionnaires and pre-test instructions were sent out by mail. After completing the questionnaire – on which respondents could also directly note any criticisms or difficulties in understanding – they were asked to complete a feedback questionnaire on various aspects of the survey. This questionnaire collected information on the time taken to complete the questionnaire, its content and length, its structure and layout, the clarity of the questions and completion instructions, and the comprehensiveness of the response categories.

Based on the results of the pre-tests for the second wave, the wording of some questions was clarified and the order of items and response categories in individual item sets was revised.

3 Population, Sample Size and Response Rate

[Population and target population] The population of the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification for the year 2018 comprises all school leavers with a higher education entrance qualification who, in the 2017/2018 school year, obtain the general higher education entrance qualification (Abitur), the subject-related higher education entrance qualification (Fachhochschulreife), or the subject- or state-specific advanced technical college certificate (Fachhochschulreife) at general education schools and vocational schools in Germany. Students who already hold a higher education entrance qualification at the time of the survey period for the first wave of the survey are excluded from the survey, as are students who do not obtain a school-leaving qualification but instead acquire eligibility for higher education entry. The survey of school leavers with a higher education entrance qualification is conducted nationwide and specifically covers the following types of school:

- Grammar schools (Gymnasium)
- Comprehensive schools/Community schools/District schools/IGS/Waldorf schools with an upper secondary level leading to the Abitur/Oberschule
- Evening grammar schools
- Specialised grammar schools (Berufliches Gymnasium)/specialised upper secondary schools (Fachoberschulen)/upper secondary level leading to the Abitur (including at OSZ, BSZ)/FOS 13/BOS 13 (AHR)
- Kollegs
- Upper vocational schools (Berufsoberschulen)/specialised upper secondary schools (Fachoberschulen)
- (Higher) vocational schools (BFS)/ higher commercial schools/ vocational school (plus)/ secondary school (Realschule) (only in Rhineland-Palatinate)/ vocational colleges (only in Baden-Württemberg)
- Trade and technical schools
- Technical academies (only in Bavaria)

[Sampling method: Quota-stratified cluster sampling] Due to the absence of, or lack of access to, lists of school leavers with higher education entrance qualifications at the individual level, it was not possible to draw a simple random sample at student level. Consequently, individuals had to be recruited via the schools. A disproportionately stratified cluster sampling procedure, as a special case of multi-stage random selection (cf. Häder, 2015, p. 169), was chosen as the sampling design. The schools or school tracks[15] constitute the 'primary sampling units' (clusters), whilst the school leavers with a higher education entrance qualification in the 2018 cohort constitute the 'secondary sampling units' within these clusters.[16] The selection population consisted of a list of all relevant schools or school tracks. Stratification was carried out on the basis of the characteristics 'federal state' and 'school type'. Within the strata, sampling rates were set for the clusters and individuals; these were based, on the one hand, on the distribution of the 2014/2015 cohort of students eligible for higher education admission according to official statistics and, on the other hand, took into account the response rates

[15] Some vocational schools comprise various tracks; however, only those leading to the higher education entrance qualification were relevant to the sample. Consequently, only selected tracks were included at certain schools.
[16] Example of the population for a cluster: all school leavers with a higher education entrance qualification from the 2018 examination cohort at a selected grammar school (Gymnasium) in a selected federal state.

from previous surveys. As part of this process, school leavers with a higher education entrance qualification from Bremen, Mecklenburg-Western Pomerania, Saxony-Anhalt and the Saarland, as well as students at evening grammar schools and vocational colleges, were assigned disproportionately high selection probabilities in order to generate sufficiently large sample sizes within these groups. Furthermore, school tracks in Hesse were selected with disproportionately high frequency, as a new requirement in the approval procedure[17] in that federal state suggested that schools' response rates would be lower.

As part of the practical implementation of the sampling procedure, clusters (schools) were initially selected in each stratum via simple random sampling. The selected schools subsequently reported their school tracks and numbers of students, enabling a second selection step to be carried out, the aim of which was to meet the sampling quotas at the individual level. In this second step, schools/school tracks were successively drawn at random from each stratum until at least 97 per cent of the stratum-specific sampling quota for school leavers with a higher education entrance qualification had been reached.[18] Following this, a complete enumeration was carried out at the schools or school tracks that had been drawn and were willing to participate, covering all students belonging to the outlined population.

Unlike the other federal states, a complete enumeration was carried out in Hamburg, meaning that the sampling method described above was not applied in Hamburg.

3.1 Wave 1

[Initial contact and address maintenance] The DZHW wrote to the selected schools and asked them to take part. In addition, the DZHW provided the schools with the criteria they were to use to identify the target individuals for the 2018 DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification (see chapter **Fehler! Verweisquelle konnte nicht gefunden werden.**).[19] As the schools were not permitted to disclose their students' contact details for data protection reasons, they merely informed the DZHW of the respective number of students. The DZHW then sent the appropriate number of survey materials for the first survey wave by mail to the relevant contact persons at the respective school branches. These contact persons were asked to pass the materials on to the students and, where possible, to allow them to complete the questionnaires as a class during lessons. The questionnaires were then ideally to be sent back to the DZHW in a single batch. However, it was also possible to complete the questionnaire at home and return it individually to the DZHW. In Mecklenburg-Western Pomerania, completing the questionnaire during school hours was generally not permitted.

In order to be able to contact those willing to participate in the second survey wave directly via the DZHW, their contact details (postal address and/or email address) were recorded on the questionnaire for the first wave. Upon receipt of a questionnaire at the DZHW, a unique identification number was stamped on both the questionnaire itself and the address section using a pagination stamp. To ensure

[17] In Hesse, it was required for the first time that schools wishing to take part in the survey had to seek permission in advance from their school council or full staff meeting. However, these meetings usually take place only once per half-year.

[18] A check as to whether the sampling quota for a stratum had been met was carried out after each sampling step. For example, if the cumulative number of students from five schools met the sampling quota for the number of students in that specific stratum, no further schools were selected after the fifth school had been chosen.

[19] At this stage, schools may in isolated cases have identified individuals as belonging to the sample who did not belong to the population (overcoverage). This occurred, for example, when school leavers had only obtained the school-based component of the advanced technical college certificate (Fachhochschulreife) without subsequently aiming for a full advanced technical college certificate (Fachhochschulreife), or when schools distributed the questionnaire to students from the wrong year group. This includes both students who have already obtained a school-based higher education entrance qualification in a previous educational programme and those who do not belong to the 2018 graduating cohort because they will not attain their higher education entrance qualification until a later date.

data protection, the address section was immediately separated from the questionnaire and, following data entry, stored and processed on a secure server separately from the survey data.

[Preliminary information] In Schleswig-Holstein, Brandenburg, Mecklenburg-Western Pomerania, Lower Saxony, Hesse, Bavaria, Saarland, Saxony and Thuringia, parental consent was required for under-age students to take part in the survey. In these federal states, from mid-November 2017 – two weeks before the survey materials were sent out – students were sent preliminary information consisting of a cover letter for the students, a letter to parents and a parental consent form. The preliminary information differed in some respects due to the data protection regulations of the respective federal states. In Hamburg, a cover letter asking students to take part in the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification was sent out as preliminary information by the Authority for Schools and Vocational Education to general education schools and by the Hamburg Institute for Vocational Education to vocational schools.

[Survey materials] The survey materials for each respondent consisted of a cover letter (including data protection information), a questionnaire, a list of academic subjects, a ballpoint pen bearing the DZHW logo, and a pre-paid envelope addressed to the DZHW, so that even if the questionnaire was completed at home, respondents could return it free of charge.

The cover letters to the students differed in some respects depending on the data protection regulations of the federal states: in addition to different cover letters, different versions of the questionnaire had to be used in four federal states (Hamburg, Hesse, Saarland, and Berlin).

- In Hamburg, question 4.7 from the standard version of the questionnaire (preferred starting salary) was moved to question 3.14 in question block 3 for reasons of space, so that respondents could be asked at the end of the questionnaire to give their consent for the survey data to be linked with school records by the data protection officer of the Authority for Schools and Vocational Training for the purposes of in-depth research questions.[20]
- In Saarland, an additional note regarding consent to participation and the use of data was included on the cover page of the questionnaire. Furthermore, a sentence in the instructions on the use of address data had to be amended (instead of 'Data analyses are always carried out anonymously', it now reads 'The data will be published in a form that does not allow conclusions to be drawn about individual persons').
- In Hesse, due to state-specific data protection regulations, no questions were asked about the respondents' parents. Specifically, the questions regarding the number of books, the parents' countries of birth and their vocational qualifications are missing (in the standard questionnaire, this information was collected in questions 5.1, 5.6, 5.7 and 5.8). Furthermore, on the last page in Block 6, there is an additional field for consent to the use of the data provided in the questionnaire. In addition, the same change to the wording was made in the instructions regarding the use of address data as in Saarland.
- In Berlin, as in Hesse, no questions were asked about the respondents' parents. Furthermore, a field for consent to the use of the data provided in the questionnaire was included before the first question. In addition, the same change to the wording in the instructions on the use of address data had to be made in Berlin as in Saarland and Hesse. Additionally, a field for consent to the use of contact details was inserted when providing the address.

In the 2018 DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification, a questionnaire was sent out for the first time to the headteachers of the selected schools. For Berlin, explicit consent to the use of the data provided was included before the first question, and a note at the end of the questionnaire advising respondents not to provide any personal data was added. For

[20] In addition, the survey documents in Hamburg were accompanied by a letter containing information on the intended data linkage. Whilst a linkage of survey data with contextual school data was not planned as part of the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018 project, it is to be implemented as part of a supplementary project.

the Saarland, an additional note on the cover page was included regarding consent to participation and the use of the data.

[Field period] The survey period for the first survey wave ran from December 2017 to October 2018.[21] Due to the contact procedure used via the schools, the DZHW was unable to directly influence the timing of the students' surveys. The reminder letter was sent to the contact persons at the schools approximately four to eight weeks after the fieldwork began. In mid-March 2018, all schools with a response rate of less than 25 per cent were reminded a second time by email.

[Measures to increase response rate] The measures to increase response rate were aimed both at providing incentives for schools to support the survey organisationally and at motivating individual respondents.

For the initial contact, the organisational support provided by the schools (and the responsible contact persons) consisted of setting aside class time for completing the questionnaires (except in Mecklenburg-Western Pomerania), collecting the questionnaires at the end of lessons and organising their return. In return, schools were offered school-specific analyses provided there was a sufficient number of responses. To encourage a high level of participation, prizes totalling ten times 300 euros were offered to the schools. However, financial incentives for schools were not permitted in Lower Saxony, Hesse, Bavaria, Saarland, Saxony and Thuringia. Measures that proved useful in increasing participation included highlighting the wider societal significance of the survey in the cover letters, ministerial calls to support the survey, and sending follow-up letters and contacting schools from which no or only a few questionnaires had been received. As a small token of appreciation for the contact persons at the schools, a notebook and a ballpoint pen were also sent.

The measure to increase response rate aimed at appealing to the individual motivation of respondents consisted of a reference to the benefits for society as a whole and for specific age groups in the cover letter to the respondents. As a material individual incentive, a laptop worth 1,300 euros, an e-book reader worth 200 euros and 20 Amazon vouchers worth 25 euros each were raffled off amongst all participating students. Students from Lower Saxony, Hesse, Bavaria, Saarland, Berlin, Saxony and Thuringia had to be excluded from the prize draw.

[Response rate] Based on the sample and survey design (see above), 2,901 schools were contacted for the first wave of the survey. Of these schools, 1,185 (41 per cent) agreed to take part and reported the number of school leavers with a higher education entrance qualification they expected to be eligible for higher education admission. This corresponded to a total of 83,988 students. The adjusted gross sample drawn from this for the first survey wave comprised 82,413 students who were contacted.[22]

[21] As the survey materials in the first wave were not sent directly to the target respondents but to the contact persons at the schools, who then distributed them, it is not possible to determine the exact date on which fieldwork began. The courier company dispatched the first parcels to the schools on 5 December 2017 (Lower Saxony, Saxony, Thuringia, Mecklenburg-Western Pomerania, Brandenburg, Rhineland-Palatinate). Some parcels sent by the DZHW itself (due to delays caused by the courier company) were dispatched as early as November. The first questionnaire returns arrived at the DZHW on 1 December (student questionnaires) and on 8 December (headteacher questionnaires). The field period was extended for as long as possible – in parallel with the other project stages. This was necessary not least because, due to increased data protection consultations and requirements, the survey could only commence later than planned in some federal states. The last questionnaire was received in October 2018.

[22] As feedback was received from the schools, some individuals were subsequently reclassified as not belonging to the sample because, in the relevant examination year of 2018, they had only completed the school-based component of the advanced technical college certificate (Fachhochschulreife). This type of *overcoverage* occurred because some schools have several academic tracks, not all of which were part of the sample but were nevertheless reported by the schools.

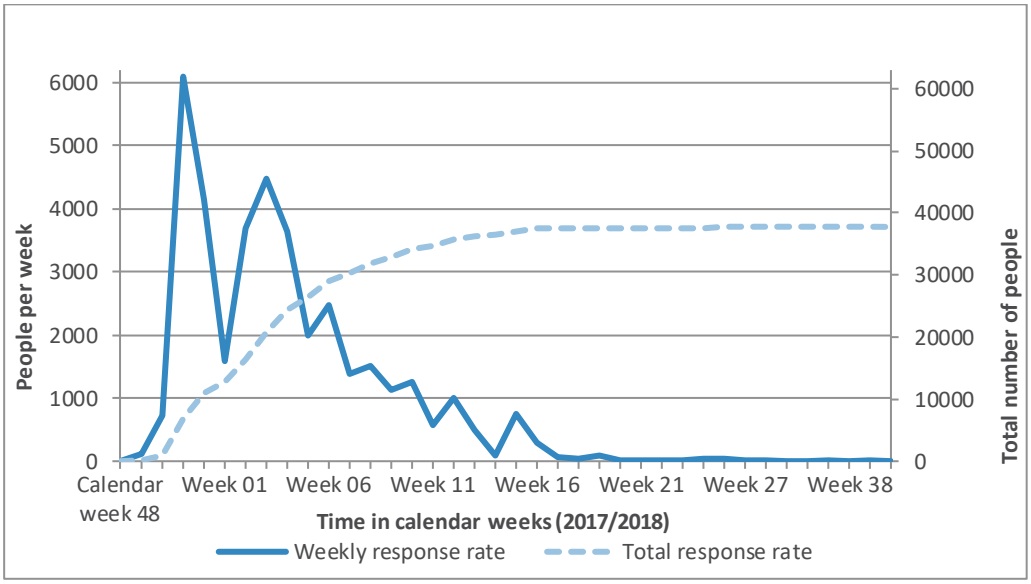
The response rate was 48.2 per cent, resulting in a net sample of 39,714 valid questionnaires.[23] Of the participants in the first wave, 63.4 per cent agreed to be contacted for further surveys and provided their address (postal and/or email).[24]

Table1: Gross and net sample sizes and response rates for the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018

	Wave 1
Gross sample	83,988
Adjusted gross sample	82,413
Net sample response rate (adjusted)	39,714
Response rate	48.2%

As shown inFigure3, the majority (87 per cent) of the questionnaires received were returned between December 2017 (calendar week 48) and February 2018 (calendar week 9). However, due to some delays in sending out invitations, a longer field period was necessary. By the end of week 17, more than 99 per cent of all valid questionnaires had been received; after that, only a few isolated questionnaires were received.

Figure3: Response rate of the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018 over time, Wave 1



[23] Eight respondents (as of 5 July 2019) who requested the subsequent deletion of their survey data were not included in the response rate. Individuals who obtained the higher education entrance qualification in a different year or who are (expected to) complete the school-based part of the higher education entrance qualification at a vocational school (BFS) were also excluded, as they do not, by definition, belong to the population. Questionnaires that were deliberately completed incorrectly or were unusable were removed from the dataset and are therefore not included in the response rate either (see chapter 4.3). However, 28 individuals who are aiming to complete the academic component of the advanced technical college certificate (Fachhochschulreife) at a type of school other than a vocational school (BFS) have been included. Finally, the response rate includes 204 individuals who do not belong to the population, although this could only be determined retrospectively (e.g. due to failing the final examination).

[24] Of these, 42 respondents from Berlin did provide an address; however, the corresponding survey data could not be used as the necessary consent was lacking (consent to the use of survey and address data was obtained separately in Berlin).

A total of 1,104 questionnaires were sent to school heads. Schools may have received several questionnaires for their heads, as several school tracks were surveyed at some schools. The response rate was 59.1 per cent, resulting in a net sample of 652 analysable questionnaires from school heads.

3.2 Wave 2

[Contact and address maintenance] In the second wave, those participants from the first wave who had provided an email or postal address for a second survey were contacted. As this was the first time the address details had been used, address updates – such as corrections to data entry errors and enquiries regarding registration details – were necessary, partly due to participants having moved in the meantime.

To analyse the effects of the survey mode on participation behaviour in a panel survey, a randomised design experiment was conducted in the second wave. As part of this experiment, participants from the first wave who had provided both a postal and an email address were randomly allocated to three experimental groups: a group surveyed by post (PAPI), an online group (computer-assisted web interview (CAWI)) and a mixed-mode group, in which participants could choose to take part either online or by post. The study participants from the first wave who had provided either a postal address or an email address only were also surveyed as part of the second wave, but were not included in the design experiment.

Around one week before the start of the field period (mid-December), all study participants who had provided a postal address received a postal notification letter to establish contact, accompanied by a leaflet containing information about the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification. Participants for whom only an email address was available received the notification letter and a PDF version of the information leaflet by email.

[Survey materials] At the start of the survey, the PAPI experimental group received a cover letter and survey materials by post, along with a pre-paid return envelope. The mixed-mode survey group, as well as those who had provided only a postal address, were also sent these survey materials, but in addition received an access code and a web link to the survey page for optional online participation. The online group, as well as respondents who had provided only an email address, received only an email covering letter containing an access code and a web link to the survey page for online participation.

Respondents from Hesse and Berlin received a version of the questionnaire which, in question block 7, collected information about their parents (number of books, parents' countries of birth and vocational qualifications), which was not permitted to be collected in these federal states during the first wave.

[Field period] The survey period for the second wave ran from late December 2018 to May 2019.

Respondents in the PAPI group were sent an initial reminder to take part in the survey (without a new questionnaire) by post approximately five weeks after the fieldwork began, and a second reminder (with a new questionnaire and a return envelope) approximately ten weeks after the fieldwork began. The mixed-mode group and those with a postal address only received an initial reminder approximately four weeks after the fieldwork began, containing a new link and access code for the online survey, and a second reminder by post approximately ten weeks after the fieldwork began, containing a questionnaire and a return envelope, as well as an access code and web link to the survey page for online participation. The online group and respondents with an email address only were first reminded online via email approximately four weeks after the start of the field period, with a renewed link and access code for the online survey and were reminded to take part in the survey in the same way approximately six weeks after the start of the field period.

Approximately 14 weeks after the start of the field phase, all respondents who had not yet taken part in the second wave and for whom an email address was available were finally reminded a third time to take part in the survey via email, with a link and access code for the online survey; this third reminder was not part of the experimental design but served solely to increase the response rate.

[Measures to increase response rates] To increase the response rate, respondents in the second wave were sent a touch-pen ballpoint pen either with the announcement letter (online group of the design experiment) or with the materials for the standardised self-administered survey. Furthermore, participation incentives worth approximately 5,000 euros were raffled off amongst the participants in the second wave.

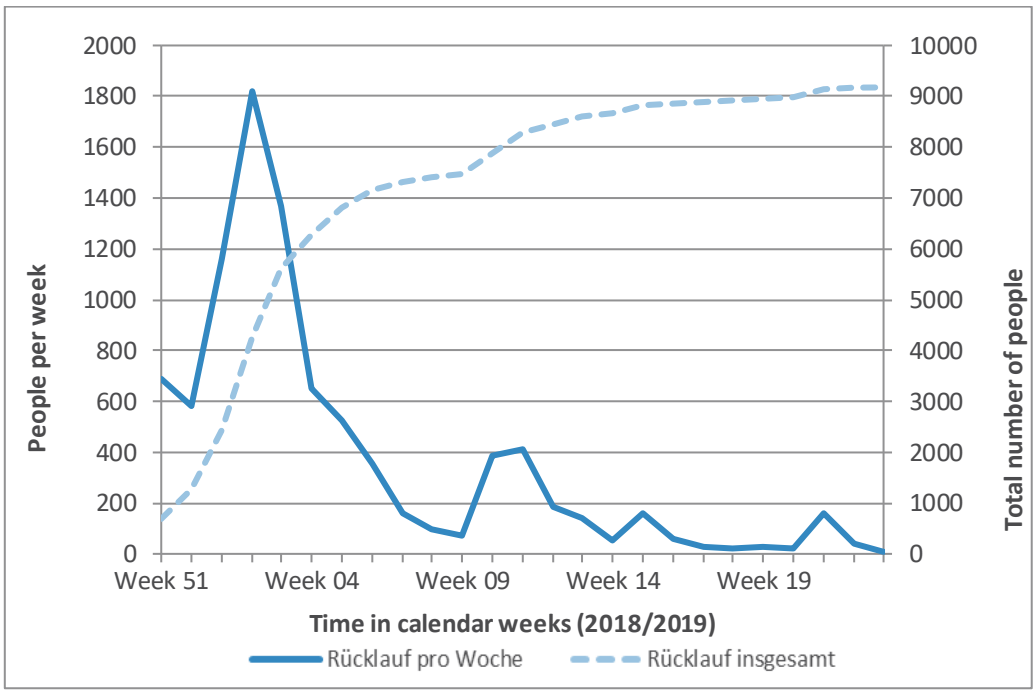
[Response rate] Based on their address details, 24,971 participants from the first wave were invited to take part in the second wave.[25] Of these, 22,364 cases could be assigned to one of the three groups in the design experiment, as both email and postal addresses were available. The group surveyed by post (PAPI) comprises 5,825 cases, the computer-assisted web interview (CAWI) 6,701 cases and the mixed-mode group 9,838 cases.[26] A further 2,235 respondents had provided only a postal address, whilst 372 respondents had provided only an email address; they were therefore not assigned to any experimental group.

A total of 10,039 people took part in the survey, 8,713 of whom completed it in full.[27] According to the definitions of the American Association for Public Opinion Research (2016), the response rate for full participation stands at 34.9 per cent [AAPOR-RR1].[28]

As shown inFigure4, over two-thirds of the responses were received within the first six weeks (calendar weeks 51 to 04). However, due to the third reminder scheduled following the design experiment, as well as delays caused by additional address verification, the field period was extended until the beginning of June.

[25] Originally, 25,175 respondents from the first wave were invited, but 204 people were subsequently excluded because they did not belong to the population. The figures reported here have already been adjusted to account for these cases.
[26] The respective allocation can be traced in the 'ausgangsgruppe' variable in the dataset.
[27] Of the respondents, a further 204 people were subsequently excluded from the panel in the second wave as they did not belong to the population (see Chapter4.3). The reported figures on invitations and participation for the second wave still include these individuals.
[28] Deviations from a simple calculation of the response rate as the ratio of survey participants to those invited may result from classification as 'non-eligible'.

Figure4: Response rate of the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018 over time, Wave 2



The DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification was affected by typical panel attrition processes in the second wave. In particular, these included a general refusal to participate in follow-up surveys (no contact details provided) and incorrect or out-of-date contact information. To compensate for this panel attrition, the non-response weights provided can be used (see chapter 5).

3.3 Wave 3a

[Initial contact and field period] The field period ran from 16 December 2019 to 26 March 2020. Around one week before the start of the field period, respondents received a letter of notification by post, containing an information leaflet and an enclosed incentive (a ballpoint pen with a touch-pen function). At the start of the fieldwork in calendar week (CW) 51 of 2018, respondents received an email containing an access code and a web link to the survey page. After approximately four and six weeks respectively, a first and second reminder was sent by email, each containing the link and access code for the online survey once again.

[Measures to increase response rates] To boost the response rate, respondents were sent a touch-pen ballpoint pen along with the introductory letter. Furthermore, incentives worth approximately 2,500 euros were raffled off amongst the participants.

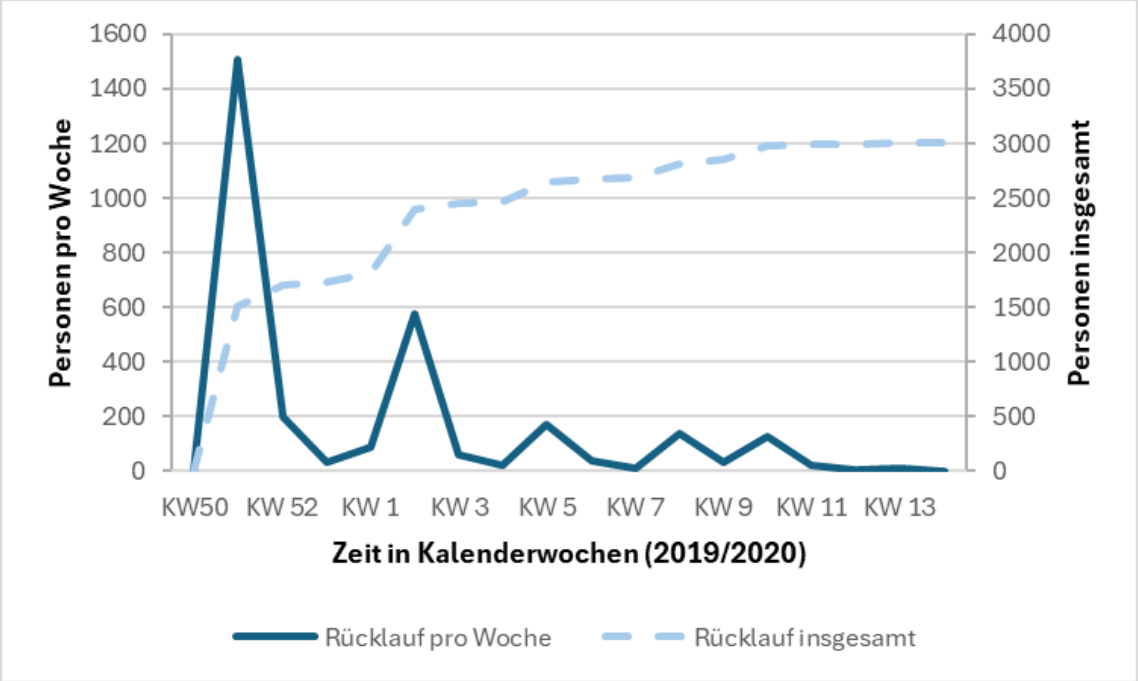
[Response rate] For Wave 3a, 4,000 participants from the second wave were randomly selected and invited, of whom 3,002 took part in the survey.[29] 2,386 of these individuals completed Wave 3a in full, whilst a further 397 completed it partially. According to the definitions of the American Association for Public Opinion Research (2016), the response rate for full participation is 59.7 per cent [AAPOR-RR1] and for partial participation 69.6 per cent [AAPOR-RR2].[30] As can be seen inFigure5, two-thirds

[29] These figures are taken from the primary research project's unadjusted project dataset. The case counts in the Scientific Use File may differ.

[30] Discrepancies compared with a simple calculation of the response rate as the ratio of survey participants to those invited may result from classification as *'non-eligible'*.

of the responses (here including all respondents and not just those who, by definition, participated at least partially) were received within the first four weeks (calendar weeks 51 to 02).

Figure5: Response rate of the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018 over time, Wave 3a



3.4 Wave 3b

For the regular survey of the third wave (Wave 3b), participants from Wave 2 and Wave 3a were invited, as well as study participants who had last taken part in Wave 1 (temporary dropouts), provided that they were willing to participate in the panel and their contact details were available. A total of N=23,161 people were therefore invited to take part in Wave 3b.

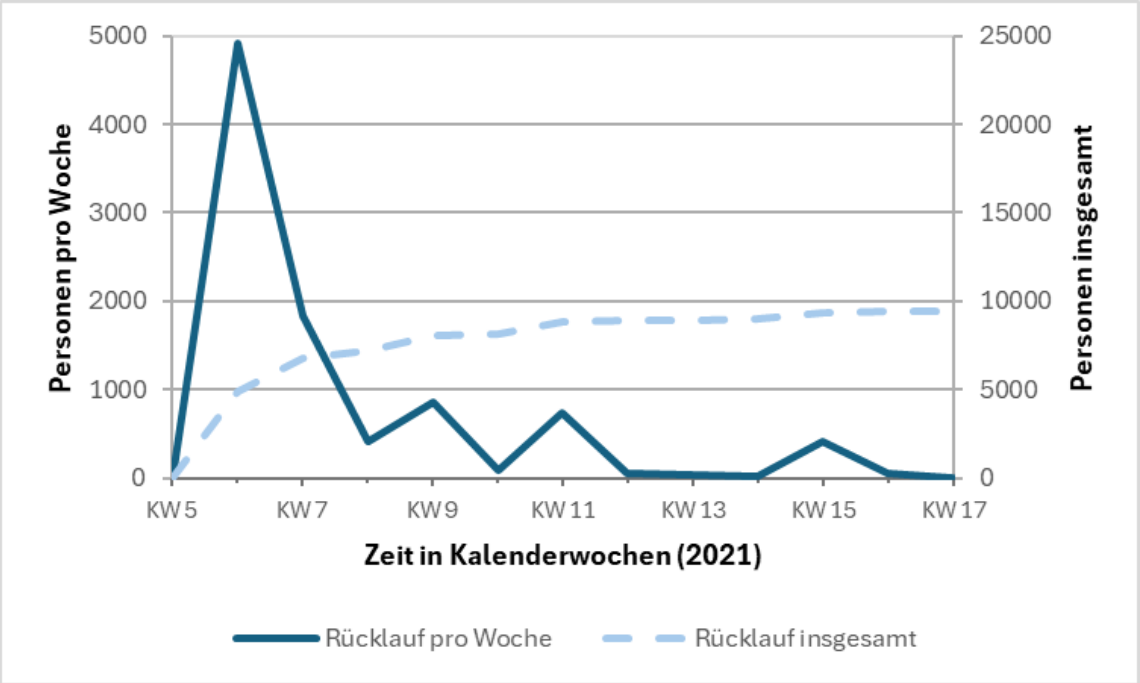
[Field period] The field period of Wave 3b (2.5 years after the baseline survey) ran from 2 February 2021 to 17 April 2021. Approximately two to three weeks before the start of fieldwork, respondents received notification letters by post, accompanied by information leaflets, provided a postal address was available. Study participants for whom only email addresses were available received the notification letters together with PDF leaflets by email. At the start of the fieldwork, study participants received an email containing an access code and a web link to the survey page. Respondents for whom only a postal address was available received the invitation by post. The first reminder was sent ten days after the fieldwork began, whilst the second and third reminders were sent three and five weeks after the start of the fieldwork respectively. A fourth reminder was sent ten days before the end of the fieldwork. Reminders were sent to all those who had not yet reached the final page of the questionnaire at that point.

[Measures to increase response rates] To boost the response rate, incentives of varying amounts were raffled off. Among respondents who took part in Wave 2, incentives totalling approximately 3,800 euros were raffled off. Respondents who had last taken part in Wave 1 (temporary dropouts) were randomly divided into two groups. Study participants in the first group were told they would be entered into a prize draw for incentives totalling approximately 1,500 euros, whilst those in the second group were entered into a prize draw for incentives totalling approximately 3,000 euros. The reason

for the different incentive amounts for these two groups was a methodological experiment designed to investigate the influence of incentive levels on increasing willingness to participate as part of the implementation study of the Student Life Cycle Panel (SLC).

[Response rate] A total of 23,161 people was invited to take part in Wave 3b. This included 3,536 respondents who had last taken part in Wave 3a and 5,287 cases who had last taken part in the second survey wave. The remaining 14,338 individuals last took part in the first wave (temporary dropouts). Of those invited, 7,424 completed the survey in full. A further 135 took part in the survey at least partially. According to the definitions of the American Association for Public Opinion Research (2016), the response rate for full participation is 32.1 per cent [AAPOR-RR1] and for partial participation 32.6 per cent [AAPOR-RR2].[31] As can be seen inFigure6, two-thirds of the responses (here including all respondents and not just those who, by definition, participated at least partially) were received within the first three weeks (calendar weeks 06 to 08).

Figure6: Response rate of the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018 over time, Wave 3b



[Panel attrition] The DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification is affected by panel attrition processes typical of such panels. Examples include the outright refusal to take part in follow-up surveys (no contact details provided) and incorrect or out-of-date contact information. To compensate for these panel attritions, the weighting variables provided can be used (see chapter 5).

[31] These are figures from the primary research project relating to the unadjusted project dataset. The case numbers in the Scientific Use File may differ.

4 Data Preparation

The steps involved in data preparation are described below. Unless otherwise stated, all steps were carried out for all surveys of the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification in 2018.

4.1 Data Transfer

The subsequent described data transfer work primarily concerned the respondents' answers from the paper questionnaires of the first and second waves. The data from the online surveys of the second and third waves were transmitted to the DZHW project managers in the primary research project by the managers responsible for the DZHW online survey software 'Zofar'. The data were transmitted in CSV format and converted to Stata DTA format for data preparation.

[Creation of a code plan] In the first and second waves, the respondents' answers from the paper questionnaires were transcribed into a machine-readable format for further processing, based on a code plan. The code plan is based on the questionnaire. In doing so, a note was made of which (sub-)question a variable was assigned to, what name this variable bore, and which numerical codes were used for the respondents' standardised answers. The variables were also numbered to determine the order of data entry. For the data transfer, the code plan, further instructions on data entry and the prepared paper questionnaires were handed over to an external service provider. The data was then entered manually by data entry clerks. For the second wave, information from the event history calendar (Question 1.11) and open responses — which were to be assigned a code by syntax — were also recorded internally for subsequent processing in the statistical software Stata.

4.2 Coding of (semi-)open Responses

The (semi-)open responses were coded (for the paper questionnaires in waves 1 and 2, this was partly carried out prior to external data entry. In this process, numerical codes were assigned to these responses using code lists. Different code lists were used depending on the variable. These comprised classification codes from official statistics (Classification of Occupations, Code List for Student and Examination Statistics, etc.) or project-specific codes already used in other studies. For some variables, new project-specific code lists were developed based on the responses found in the data for the 2018 cohort of students eligible for higher education. For a semi-open-ended question, new variables with numerical codes were created only if the responses could not be assigned to the existing (closed) response categories. The coding of subjects, higher education institutions and occupations was syntax-based.

It should be noted that the coding of parents' occupational activity was carried out by the IEA Hamburg organisation in accordance with the 2010 Classification of Occupations. In addition to the openly collected occupational data on occupational title and main area of activity, the stated occupational position and, in individual cases, further contextual information, such as the parents' educational qualifications, were also taken into account.

Table2 documents the coded characteristics and the respective code lists used. The values of the individual variables are documented in the variable report. The dataset contains only coded numerical variables; the open-ended responses themselves are not included in the dataset.

Table2: Coded variables and code lists in the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018

Characteristic	Code list
Occupational titles	German Classification of Occupations 2010 (KldB 2010) (5-digit)[a] International Standard Classification of Occupations 2008 (ISCO-08)[b] International Socio-Economic Index of Occupational Status (ISEI)[c] Occupational field definition (BIBB 2010)[d]
Countries	Project-specific coding Destatis BEV code[e] ISO 3166-1[f]
Languages	Project-specific coding
Subjects/degree programmes	Destatis Code List for Student and Examination Statistics (Winter Semester 2017/18)[g] supplemented by project-specific coding
Higher education institutions	German higher education institutions: Destatis code list for student and examination statistics (winter semester 2017/18)[g] (supplemented by project-specific coding) Foreign higher education institutions: Project-specific coding by country of the higher education institution
Other open queries	Project-specific coding or assignment to predefined categories

a https://statistik.arbeitsagentur.de/DE/Static-Content/Basics/Classifications/Classification-of-Occupations/KldB2010-first-version/Generic-Publications/Classification-Systems/KldB2010-first-version-Systematic-Directory.xlsx (accessed on 9 April 2026)
b https://www.ilo.org/publications/international-standard-classification-occupations-2008-isco-08-structure (accessed on 9 April 2026)
c http://www.harryganzeboom.nl/isco08/qa-isei-08.htm (accessed on 9 April 2026)
d https://metadaten.bibb.de/download/pdf/75_20160113_1452671666/Berufsfeld-Definitionen%20des%20BIBB (accessed on 9 April 2026)
e https://www.destatis.de/DE/Methoden/Klassifikationen/Staat-Gebietsystematik/Staatsangehoerigkeitsgebietsschluessel_pdf.pdf?__blob=publicationFile&v=27 (accessed on 9 April 2026)
f https://www.iso.org/iso-3166-country-codes.html (accessed on 9 April 2026)
g https://www.statistik-nord.de/fileadmin/Dokumente/Surveys/Hochschulstatistiken/1_Studenten_2_Pr%C3%BCfungen/Schluesselverzeichnis_WS_2017_Stand_2018_01_04.xlsx (accessed on 9 April 2026)

4.3 Data Checking and Cleaning

[Preliminary manual correction of paper questionnaires in Waves 1 and 2] Even before the data was entered (see chapter 4.1), the paper questionnaires were manually checked and, where necessary, the respondents' answers were adjusted.[32] This was primarily intended to facilitate data entry. To this end, the format of the existing entries was altered as a first step. For example, entries that were difficult to read or crossings-out made by respondents were clarified; numerical values were entered right-aligned into the boxes provided for this purpose; and verbal descriptions of grades were converted into numerical values (e.g. 'good' = 2.0). In cases where multiple responses were given to questions that required a single selection, additional codes were sometimes added.

[Software-assisted correction of paper questionnaires in Waves 1 and 2] Following data transfer, a comprehensive check and correction of the data was carried out using DZHW's own software and the

[32] The number of corrections made was not recorded centrally, but only on the paper questionnaires, and therefore cannot be systematically reconstructed.

statistical software Stata. The aim was, on the one hand, to identify any errors that may have occurred during the preliminary manual correction and data transfer, and, on the other hand, to identify inconsistent responses from respondents.

For this purpose, the questionnaire data collected was imported into a database. Valid value ranges and combinations of answers were then defined and checked on the basis of formal rules. The following types of checks were carried out:

- *Test of value ranges:* It was checked whether the recorded value for a variable fell within the value range defined for that variable. For example, for variables on a 5-point scale, only values from 1 to 5 were permitted; when coding the open-ended occupational details, the codes assigned had to be 5 digits long. Numerical values in open responses were treated differently depending on the variable: values that were highly unlikely but theoretically possible (e.g. 90 working hours per week) were not altered. Values that were obviously impossible (e.g. 70 sick days in four weeks) were classified as implausible and coded accordingly as missing values.
- *Test of adherence to filter procedures:* Based on the questionnaire's defined filter procedure, checks were carried out to determine, firstly, whether information should have been provided for the respondent in question but was missing (completeness check), and, secondly, whether information was provided for the respondent that should not have been (filter violations).
- *Test of variable consistency:* The consistency of the information within the questionnaire was checked.

In total, several hundred consistency rules were defined and tested. In the case of missing, incorrect or implausible values, the paper questionnaire was first consulted to check whether the relevant value had been incorrectly (or not at all) transcribed. If this was not the case, adjustments or corrections were made on the basis of rules specifically defined for the individual questions, as well as other information in the questionnaire. The relevant rules and coding instructions for dealing with inconsistencies were recorded centrally. In cases of doubt, a specific missing value code was assigned (see chapter 0). Error corrections were partly documented on the questionnaire[33] and were generally checked by at least one other person. Error corrections carried out as part of the software-assisted verification using Stata were documented as a Do-file and are reproducible. The corrections were made in duplicates of the original variables. In these cases, the published dataset contains only the validated variables.

[Deletion of cases] Paper questionnaires deliberately completed incorrectly by respondents were removed from the dataset, as were cases that did not belong to the population (e.g. incorrect year of birth, or only the school-based component of the advanced technical college certificate (Fachhochschulreife); see chapter **Fehler! Verweisquelle konnte nicht gefunden werden.**). In particular, 204 cases were subsequently deleted in the second wave; these did not belong to the population of school leavers with higher education entrance qualifications, largely because they had failed their examinations. Furthermore, cases were removed from the dataset that had been received twice, as respondents (in the second wave) had completed both the paper questionnaire and the online questionnaire. As a rule, the questionnaire received later was deleted, unless the first questionnaire had been not fully completed. Finally, all cases received after the end of the defined field period were deleted. In the third survey wave, there were a total of 29 temporary dropouts (who had not participated in the second wave), for whom it became apparent during survey 3b of the third wave that they did not belong to the population of school leavers with a higher education entrance qualification, as they had only obtained the school-based component of the advanced technical college certificate (Fachhochschulreife) in the 2017/18 academic year.

[33] The documentation of error corrections was made by hand on the paper questionnaires and therefore cannot be systematically reconstructed.

4.4 Generation of Variables

In addition to the variables containing the respondents' coded answers, the dataset also includes generated variables. These consist, on the one hand of variables with numerical codes derived from originally open responses (see chapter 4.2). On the other hand, some variables were modified for data protection reasons (see chapter 6) and variables frequently required in the research field were generated from the values of a source variable (e.g. aggregation of subjects into areas of study and subject groups, or derivation of higher education institution type and location from the higher education institution variables). The vast majority of the generated variables were already created by the primary research project. An overview of all variables generated for the 2018 DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification, as well as detailed documentation of the individual variables specifying their respective values and calculation rules, can be found in the dataset report and on the FDZ-DZHW search portal.[34]

4.5 Coding of Missing Values

A system for coding missing values was developed for the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification in line with the project-specific requirements.Table3 illustrates the missing value coding system used.

Table3: System of missing values in the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018

Value	Description	Explanation
-999	Don't know	'Don't know', explicitly stated
-998	No answer	The respondent did not provide an answer. This code is also assigned if a respondent has not answered an item set at all (which is normally coded as 0 for 'not mentioned' and 1 for 'mentioned').
-996	Interview break-off	Assigned to all subsequent items/questions if no further information is available from a certain point onwards.
-995	Did not participate (panel)	The respondent did not take part in the relevant wave.
-989	Filtered	Missing due to filtering (see flowchart)
-988	Does not apply	This code is not assigned during filter procedure, but rather when a 'does not apply' response category is explicitly provided or a response category has been excluded by other response categories.
-987	Missing by design (Questionnaire split)	The question was not asked due to a split design or specifications, e.g. requirements set by the federal states for a sub-group of respondents.
-986	Missing by design (wave)	The question was not asked in this wave.
-968	Implausible value	Defined differently depending on the question (e.g. value range, implausible multiple responses)
-967	Anonymised	Assigned as part of the anonymisation process (see chapter 6)
-966	Not determinable	Assigned to open responses or comments that could not be coded or validated.
-949	Not valid	Invalid (e.g. obscured or crossed out)
-929	Loss of data	Data not recorded or not recorded correctly, e.g. due to technical problems.

[34] https://metadata.fdz.dzhw.eu/en/start

4.6 Generation of the Dataset

[Merging of waves] This data package comprises the following datasets:

Individual data:

- Survey data from students/school leavers with a higher education entrance qualification (first, second and third waves)
- Survey data from headteachers (first wave)

Spell data set:

- Activity episodes from the event history calendar (second wave)
- School episodes from the episode module (third wave)
- Vocational Training episodes from the episode module (third wave)
- Study episodes from the episode module (third wave)
- Work placment (internship) episodes from the episode module (third wave)
- Further education episodes from the episode module (third wave)
- Employment episodes as an employee from the episode module (third wave)
- Episodes of unemployment from the episode module (third wave)
- Self-employment episodes from the episode module (third wave)

The order of the variables in the datasets follows the order of the corresponding questions in the respective survey instrument. Variables that were collected across several waves (panel variables) are listed in the order in which they were first collected.

[Creation of the dataset] The data provided by students and school leavers with a higher education entrance qualification in the four surveys comprising the first to third waves were merged into a single dataset. This is in 'long' format, i.e. for each respondent (student), there are four data rows containing the information from the respective wave (distinguishable via the 'wave' variable). For respondents who took part in only some of the waves, the rows corresponding to the waves in which they did not participate contain the appropriate missing codes ('no participation (panel)'). Variables that were collected only in individual surveys are marked as missing for the other waves ('missing due to survey design (wave)'). Variables generated additionally during data preparation (see chapters 4.4, 5 and 6) have been included.

The dataset containing the information provided by school heads includes both a randomised identification number for the respective school ('schul_id') and for the school management(s) surveyed ('schulleitung_id'), as several school tracks were surveyed at some schools. These two variables enable the school heads responses to be linked to the survey data from students/school leavers with a higher education entrance qualification. However, information from school heads is not available for every student/school leaver with a higher education entrance qualification.

The dataset containing activity episodes is based on the responses given by students/school leavers with a higher education entrance qualification in the second wave to Question 1.11 regarding activities undertaken since January 2014. The dataset is in 'long' format, i.e. for each activity mentioned, there is a data row containing the type, start and end dates of the respective episode. The data is linked to the survey data via the respective identification number of the school leavers with a higher education entrance qualification ('pid').

The episode data from the episode module of the two third-wave surveys were processed by the primary research project in such a way that each loop or split episode is treated as a separate element and subsequently merged into a single dataset. As part of the creation of the Scientific Use File, these were – for greater clarity – split into sub-datasets, each of which contains a specific episode type and the associated detailed information from the episode module. The datasets from the episode module contain both cross-dataset variables (e.g. start/end date of the episode), which are present in all sub-datasets, and the variables specific to each episode type (e.g. all variables relating to the details of the work placement (internship) episode), which are only present in the relevant sub-dataset. The datasets from the episode module can be combined by data users (e.g. using the Stata command 'append') to create a single spell data set.

[Data format] All datasets are provided in Stata and SPSS formats as standard.

4.7 Assignment of Variable Names, Variable Labels and Value Labels (across Waves)

[Assignment of variable and value labels] For variable and value labels, wording from the questionnaire was adopted or concise abbreviations of this wording were chosen. The variable labels are generally based on the corresponding question text. Depending on the question type, the value labels are based on the texts of the response categories or a combination of the texts from the question and the response categories. The labels are available in German and English.

[Variable naming] Variables were named according to content-based criteria; in other words, the names are derived from the content of the questions. For indicators used across multiple survey waves, the names of the associated variables were harmonised by assigning them an identical stem. For question batteries, a common variable stem was chosen and the individual items were numbered (e.g. unterricht01–unterricht10). Variables generated through aggregation or derivation from another variable adopted (where possible) the name of the original variable, supplemented by a suffix to identify the derivation or aggregation (e.g. '_isco'), and were placed after the respective source variable.

Detailed information on variable naming in the 2018 DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification can be found in the dataset report on the FDZ-DZHW search portal.[35]

[35] https://metadata.fdz.dzhw.eu/en/start

5 Weighting

The weighting of the data serves to compensate for discrepancies between the sample and the defined target population. The design weight is also necessary to account for the disproportionate nature of the sample. First, a general introduction to the procedure and a description of the weights constructed is provided. The weighting procedure is then described.

5.1 Procedure and Application Notes

[Causes of sample bias] Two processes are key to the biases in the sample:

- **Design-related bias:** Disproportionalities were deliberately introduced during the sampling process in order to increase the number of cases in certain relevant subgroups.
- **Non-response bias:** Missing data processes (e.g. non-participation, inability to contact respondents, loss in the post) lead to a reduced response rate and thus to a difference between the gross and net sample. If these processes are unsystematic (Missing Completely at Random), they can be ignored.[36] However, they are usually subject to a systematic process (Missing at Random, Missing Not at Random), which requires modelling.[37]

[Conceptual approach] Both a design weight and two cross-sectional weights are provided for the analyses of the first wave. For the second wave, two longitudinal weights were created. In the case of randomised selection procedures, the design weight is derived directly from the sampling design. To generate the cross-sectional weights, as in previous surveys of prospective higher education students, cell weighting was used to adjust the data to selected characteristics of the population.

Two cross-sectional weights are provided for the analysis of the third-wave surveys. For each survey, there is one weight that takes into account temporary dropouts in the second wave and another in which these cases are not taken into account.

The weights provided in the dataset for the respective waves are shown in the followingTable4 .

Table4: Weights provided for the DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification 2018

Variable name	Description
desgew	Design weight for Wave 1
gew1	Cross-sectional weight for wave 1 (for state-level analyses)
gew2	Cross-sectional weight for Wave 1 (for national analyses)
gewa	Longitudinal weight for Wave 2 (for state-level analyses)
gewb	Longitudinal weight for Wave 2 (for national analyses)
gew183a	Cross-sectional weight for Wave 3a (including temporary dropouts in Wave 2)
gew183a_ohne_ausfall	Cross-sectional weight for Wave 3a (excluding temporary dropouts)
gew183b	Cross-sectional weight for batch 3b (including temporary dropouts in Wave 2)
gew183b_ohne_ausfall	Cross-sectional weight for wave 3b (excluding temporary dropouts)

[36] Insofar as the loss of statistical power resulting from the reduction in sample size is considered irrelevant.
[37] See Rubin (1987) for a discussion of the different forms of missing data processes.

[Notes on the application of weights] The weights generated are *probability weights*, which can be taken into account in Stata using specific options.[38] It is essential to note that weights only serve as meaningful correction factors for those analyses in which the variables used for weighting are included in the analytical model or are related to it. For this reason, weights should always be used with caution and with a focus on the research question being analysed. The procedure for generating the weights for the dataset is described in more detail below.

5.2 Weighting the Dataset

The following section describes the weighting separately for the first two waves and the two surveys of the third wave.

5.2.1 Wave 1

[Design weight] In the present study design, the design weight is calculated directly from the sampling plan. A specific design weight is determined for each stratum, with the strata being defined on the basis of the characteristics of federal state and school type[39] (see chapter **Fehler! Verweisquelle konnte nicht gefunden werden.**). The following information from the sampling plan is required for each stratum:

- Number of schools or school tracks in the population
- Number of schools or school types sampled per stratum

Using this information, the specific selection probability π is determined for each stratum. This is calculated using the formula:

$$\pi = \frac{\mathrm{n(sample)}}{\mathrm{N(population)}}$$

The design weight is ultimately calculated as the reciprocal of the selection probability: $1/\pi$.

[Cross-sectional weighting] For the first wave, cross-sectional weights were created using cell weighting. In this process, the data were adjusted to the population on the basis of the characteristics of gender, school groups[40] and federal state, in order to compensate for any biases in the net sample compared with the distribution in the population with regard to these characteristics.[41] These non-response biases may result from both unit and item non-response, as well as from the disproportionate sample design. Unlike the design weights, the cross-sectional weights for the first wave (gew1 and gew2) do not include all respondents from the full survey in Hamburg. Instead, a subsample of schools/school tracks was drawn at random in Hamburg in accordance with the population and included in the cross-sectional weights.

For each cell in the weighting matrix, the proportion that this cell represents in the population is first determined on the basis of the target figures (pN). A similar ratio is also calculated for the net sample

[38] See the Stata help (command: *help weights*) for further details.

[39] These are the nine types of schools that were also decisive for the sampling (see chapter **Fehler! Verweisquelle konnte nicht gefunden werden.**).

[40] These were the following groups: (1) Gymnasium/comprehensive school/Waldorf school, (2) evening grammar school/Kolleg, (3) specialised grammar school (Berufliches Gymnasium)/other vocational schools leading to the general higher education entrance qualification (Abitur), (4) specialised and upper vocational schools (leading to the FHR), (5) other vocational schools (leading to the FHR qualification). The classification into school groups is based on the respondents' statements regarding the school or school stream they attended.

[41] The data was based on the figures available at the time of weighting (December 2018) for school leavers with higher education entrance qualifications from the 2017/18 academic year. These were taken from Specialist Series 11 (Series 1 for general education schools and Series 2 for vocational schools). However, this statistical series only contains the graduate figures for 2017.

(pn), which indicates the proportion of realised cases in the respective cell divided by the total number of cases in the sample. The weighting factor is obtained by calculating the quotient of these two proportion values (pN/pn).

This procedure is carried out twice. For the first weight 'gew1', the weighting factors are calculated separately for each federal state. The reference group for determining the proportion values pN and pn is then the number of cases in the population or sample in the respective state. The weight 'gew2' additionally takes into account the relationships between the states, as the proportion values here are determined in relation to the nationwide case numbers.

[Normalisation to the sample size] As is standard practice in social science research, the calculated design weight was normalised to the sample size. As the calculation of the cross-sectional weights is based on relative frequencies, normalisation to the sample size has already taken place.

[Weight trimming] No weight trimming was carried out for the design weight. In the case of cell weighting, cells were merged in some instances. Cell merging was absolutely necessary whenever a cell contained cases according to official statistics, but no corresponding cases were present in the sample. Such 'empty cells' were then merged with another cell according to substantive criteria.

Furthermore, extreme weighting factors were checked and, in individual cases, also merged with other cells. In this context, all weights with values less than 0.3 and greater than 6 were checked. However, cell merging was only carried out in the event of significant deviations from the threshold values and where a substantive merger with another cell was possible. In all cases, cell merging was carried out only between school types. For content-related reasons, cells were not merged across federal states or between men and women.

5.2.2 Waves 2, 3a and 3b

[Longitudinal weighting] For the second wave and all surveys in the third wave, non-response weights were calculated using propensity scores.[42] This method takes advantage of the fact that, unlike in non-response processes between the gross and net samples in the first wave, a wealth of information on the respondents from the respective previous waves is already available in the event of non-response in subsequent waves. This information is used as covariates in a logistic regression aimed at predicting the probability of participation in the subsequent wave. The probability that a person i will remain $Pr(Yi = 1)$ in the panel can then be calculated as

$$Pr(Yi = 1 \mid Xi) = \beta 0 + \beta 1 Xi1 + \ldots + \beta p Xip$$

Where $Xi = (Xi1, Xi2 \ldots Xip)$ represents the list of relevant predictors from the previous wave. The weighting factor for each person is then calculated as the inverse of their individual probability of remaining in the panel. Respondents with a high probability of remaining are assigned a low weighting factor, whilst those with a low probability of remaining are assigned a high weighting factor.

The quality of the calculated weights depends largely on the specification of the estimated logistic regression model and, in particular, on the selection of relevant predictors (see Chen et al., 2014, p. 3; Little & Vartivarian, 2003). To take this into account, an exploratory, multi-stage selection procedure was used: in a first step, all variables for which data were not available for all respondents due to the questionnaire's filter procedure were removed. In a second step, bivariate logistic regressions on the probability of retention in Wave 3 were calculated for all remaining variables. Variables that showed a significant effect ($p < 0.05$) and a pseudo-R^2 of more than 0.01 were finally included in the logistic model.[43] Variables that did not contribute to improving the estimation model when all other

[42] The methodology is based on the logic of propensity score matching, which dates back to Rosenbaum and Rubin (1983) (see also Blumenstiel & Gummer, 2015).

[43] The pseudo-R^2 proposed by Tjur (2009) was used to evaluate logistic regressions (see also: Cramer, 1999).

covariates were controlled for were removed again. As it is known from comparable panel studies (e.g. the SOEP) that, in addition to survey content data, information about the fieldwork can also be an important predictor of the probability of continued participation (Schupp, 2004), such information was utilised and separate models were estimated for the various survey and incentive groups (see sections 1 and 3). In order to estimate the probability of participation for each individual, missing values for all variables were incorporated into the model as an additional category. For metric variables, missing values were set to 0 and an additional dummy variable was included in the model for each variable, indicating whether a case had a missing value for the corresponding metric variable. The key criterion for modelling the default processes was to find models with the greatest possible explanatory power whilst remaining parsimonious, thereby keeping the increase in variance as low as possible.

A further difficulty with propensity score matching for calculating non-response weights is that respondents may have a very low estimated probability of participation and, as a result, be assigned particularly high weights, which can lead to high variance in weighted estimates of population proportions. A common approach is to trim particularly large weights (Potter, 1990). The method is based on the assumption that weights follow a beta distribution. All weights above the 99th percentile were truncated at this threshold. Following weighting, the non-response weight was normalised to the sample size to avoid creating discrepancies in the sample size in subsequent analyses.

The final cross-sectional weights are ultimately calculated by multiplying the non-response weight by the calibrated design weight from the first wave. This serves to compensate for differences in the probability of being selected for the first-wave survey and to adjust the characteristics of gender, school groups and federal state to the population. In this case, the trimming and standardisation of the resulting longitudinal weight were carried out only after the multiplication.

The final cross-sectional weights for the second wave are available for analyses at state and national level (‘gewa’ for state-level analyses, ‘gewb’ for national analyses). For each sub-survey of the third wave, the cross-sectional weights were calculated both with and without temporary dropout, in each case to ensure comparability with the same logit model.

6 Anonymisation

[Data protection framework] Personal data[44] collected in voluntary surveys conducted by the DZHW is subject to the EU General Data Protection Regulation (GDPR) and the Federal Data Protection Act (BDSG) in its revised version of 30 June 2017.[45] Accordingly, personal data intended for disclosure for secondary scientific use (in the absence of a declaration of consent for the secondary use of the personal data) must generally be processed in such a way that "the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is stored separately and is subject to technical and organisational measures ensuring that the personal data cannot be attributed to an identified or identifiable natural person" (Article 4(5) of the General Data Protection Regulation (GDPR); see also Article 89 of the GDPR and Recital 26 of the GDPR).

[Ensuring data protection at the Research Data Centre for Higher Education Research and Science Studies (FDZ-DZHW)] At the Research Data Centre for Higher Education Research and Science Studies (FDZ-DZHW), the data protection of respondents is ensured through legal and organisational measures, technical measures and statistical measures[46], which can be combined in various ways. These combinations result in four potential modes of access, which are summarised in the tableTable5 .[47]

Table5: Modes of access at the Research Data Centre for Higher Education Research and Science Studies (FDZ-DZHW)

	Campus Use File (CUF): Download	Scientific Use File (SUF): Download	Scientific Use File (SUF): Remote Desktop	Scientific Use File (SUF): On-site
Legal and organisational measures	very low	low	moderate	high
Technical measures	very low	low	moderate	high
Statistical Measures	very high	high	moderate	low

The more strictly data access is controlled and restricted from a legal, organisational and technical perspective, the lower the risk of data re-identification. Consequently, there is less need to perform

[44] "'Personal data' means any information relating to an identified or identifiable natural person (hereinafter referred to as the 'data subject'); a natural person is regarded as identifiable if they can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person' (Article 4(1) of the General Data Protection Regulation (GDPR)).

[45] The General Data Protection Regulation (GDPR) applies in principle within the EU and therefore also to the DZHW. The Federal Data Protection Act, as amended on 30 June 2017 (Act on the Adaptation of Data Protection Law to Regulation (EU) 2016/679 and on the Implementation of Directive (EU) 2016/680 (EU Data Protection Adaptation and Implementation Act – DSAnpUG-EU)) applies in part in addition to the GDPR, as DZHW GmbH is legally regarded as a federal public body (Section 2(3) of the Federal Data Protection Act). The Federal Government holds an absolute majority of the shares in DZHW GmbH, and the department performs tasks of federal public administration in the broadest sense.

[46] The FDZ's data protection framework is based on the portfolio approach developed by Lane, Heus and Mulcahy (2008, p. 6 ff.), which has already been adopted by the Leibniz Institute for Educational Trajectories (LIfBi) (cf. Koberg, 2016, p. 699 ff.) and the FDZ of the Federal Employment Agency within the Institute for Labour Market and Employment Research (see Hochfellner, Müller, Schmucker, & Roß, 2012, p. 9 ff.).

[47] For further details, see https://www.fdz.dzhw.eu/en/data-usage

anonymisation – i.e. to reduce the amount of information the data contain – using statistical measures; in other words, its analysis potential remains greater.

For the data from the 2018 DZHW Panel Study of School Leavers with a Higher Education Entrance Qualification, the only modes of access available are 'Scientific Use File (SUF): Remote Desktop' and 'Scientific Use File (SUF): On-Site'. The specific statistical anonymisation measures carried out are explained in more detail below.

[Statistical anonymisation measures] As part of the statistical anonymisation measures, all information was first checked to determine whether it could be used to directly identify individuals. These so-called *direct identifiers*, such as names, addresses and email addresses, were already recorded in a separate dataset during the field period and are therefore not included in the dataset. The original identification number was removed and replaced with a new, randomly assigned identification number.

Subsequently, *quasi-identifiers* were identified – that is, information which, in combination or by reference to external information, is capable of indirectly identifying an individual. To prevent the data from being uniquely traced back to individuals, these key characteristics were either aggregated or withheld entirely, depending on the mode of access. Variables newly generated during the aggregation process were marked in the dataset by the suffix '_g#' in the variable name and placed after the respective source variable.

Furthermore, Ebel and Meyermann recommend deleting open responses "even if the questions in doubt are unproblematic in themselves. This is because there is a risk that study participants may have disclosed sensitive information in response to questions with open response categories – questions that are in fact harmless – which could lead to their identification" (Ebel & Meyermann, 2015, p. 5). Accordingly, no open responses are released.

Finally, a check was carried out to determine whether the data contained sensitive information – for example, regarding health, sexual orientation and political views – for which no additional consent had been obtained from the respondents for secondary use. Whilst such information is not necessarily suitable for data re-identification of individuals or institutions, it could prove useful in the event of data re-identification (see Koberg, 2016, p. 694) and is therefore particularly worthy of protection (Art. 9 General Data Protection Regulation (GDPR), Recital 51 GDPR). For sensitive information too – depending on the mode of access – it was determined whether it may only be released in aggregated form, if at all, or not at all.

An overview of which variables are contained in the released data, how they may have been aggregated and via which mode of access they can be used can be found in the respective dataset report for the datasets in the data package, as well as in the FDZ-DZHW search portal.

7 References

Blumenstiel, J. E. & Gummer, T. (2015). *Prävention, Korrektur oder beides? Drei Wege zur Reduzierung von Nonresponse Bias mit Propensity Scores.* In J. Schupp & C. Wolf (Eds.), Nonresponse Bias. Qualitätssicherung wissenschaftlicher Umfragen (pp. 13-44). Wiesbaden: Springer Fachmedien.

Chen, Q., Gelman, A., Tracy, M., Horris, F. H., & Galea, S. (2014). Incorporating the sampling design in weighting adjustements for panel attrition. *Statistics in Medicine*.

Cramer, J. S. (1999). Predictive performance of the binary logit model in unbalanced samples. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *48*(1), 85-94.

Ebel, T. & Meyermann, A. (2015). *Hinweise zur Anonymisierung von quantitativen Daten.* Forschungsdaten Bildung informiert. Bd.3. Frankfurt am Main: Deutsches Institut für Internationale Pädagogische Forschung.

Hochfellner, D., Müller, D., Schmucker, A., & Roß, E. (2012). *FDZ-Methodenreport. Datenschutz am Forschungsdatenzentrum (Nr. 06)*. Nürnberg: Institut für Arbeitsmarkt- und Berufsforschung (IAB).

Koberg, T. (2016). Disclosing the National Educational Panel Study. In H.-P. Blossfeld, J. v. Maurice, M. Bayer, & J. Skopek (Eds.), *Methodological Issues of Longitudinal Surveys. The example of the National Educational Panel Study* (pp. 691-708). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-11994-2_38

Lane, J., Heus, P., & Mulcahy, T. (2008). Data access in a cyber world: Making use of cyberinfrastructure. *Transactions on Data Privacy, 1*(1), 2-16.

Little, R. J. A. & Vartivarian, S. (2003). On weighting the rates in non-response weights. *Statistics in Medicine 22*, 1589-1599.

Potter, F. J. (1990). A study of procedures to identify and trim extreme sampling weights. *Proceedings of the Survey Research Methods Section*, 225-230.

Rosenbaum, P. R. & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41-55.

Rubin, D. B. (1976). Inference and missing data. *Biometrika, 63*(2), 581-592.

Schupp, J. (2004). Gewichtung in der Umfragepraxis. Das Beispiel SOEP. Available online at: http://eswf.uni-koeln.de/lehre/04/04_05/schupp.pdf

Tjur, T. (2009). Coefficients of determination in logistic regression models - A new proposal: The coefficient of discrimination. *The American Statistician 63*, 366-372.

Appendix

Table6: Documentation of the origin of reused questions in Wave 1 (student survey)

Question number	Original (O)/ Modified (M)	Citation
1.11	M	Centers for Disease Control and Prevention (CDC) (1993). Behavioral Risk Factor Surveillance System Survey Questionnaire. Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, p. 3. https://www.cdc.gov/brfss/questionnaires/pdf-ques/93brfss.pdf (accessed on 10 November 2022). FDZ-LIfBi (2022). Codebook. NEPS Startkohorte 5 — Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 16.0.0, p. 1554. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/16-0-0/SC5_16-0-0_Codebook_de.pdf (accessed on 10 November 2022).
1.12	M	OECD (2017). PISA 2015 Technical Report, OECD Publishing, pp. 315-317. https://www.oecd.org/pisa/data/2015-technical-report/PISA2015_TechRep_Final.pdf (accessed on 10 November 2022). Mang, J., Ustjanzew, N., Leßke, I., Schiepe-Tiska, A. & Reiss, K. (2019). PISA 2015 Skalenhandbuch. Dokumentation der Erhebungsinstrumente. Münster: Waxmann. pp. 51-52. https://www.pisa.tum.de/fileadmin/w00bgi/www/Berichtsbaende_und_Zusammenfassungungen/Skalenhandbuch_PISA2015_openaccess.pdf (accessed on 10 November 2022).
1.15	O	Beierlein, C., Kovaleva, A., Kemper, C. J. & Rammstedt, B. (2014). Allgemeine Selbstwirksamkeit Kurzskala (ASKU). Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS). doi:10.6102/zis35.
1.16	M	Items 1, 2, 7: Adapted from FDZ-LIfBi (2022). Codebook. NEPS Startkohorte 5 — Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 16.0.0, pp. 730, 729, 728. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/16-0-0/SC5_16-0-0_Codebook_de.pdf (accessed on 10 November 2022). Items 3, 4, 5, 6, 8, 9: Adapted from Stocké, V. (2005). Einstellung zu Bildung. Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS). doi:10.6102/zis136.
2.1	M	Adapted from FDZ-LIfBi (2021). Codebook NEPS Startkohorte 4 — Klasse 9. Schule und Ausbildung. Scientific Use File Version 12.0.0, pp. 6971-6973. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC4/12-0-0/SC4_12-0-0_Codebook_de.pdf (accessed on 10 November 2022).
3.8	M	Adapted from FDZ-LIfBi (2021). Codebook NEPS Startkohorte 4 — Klasse 9. Schule und Ausbildung. Scientific Use File Version 12.0.0, pp. 6713-6720. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC4/12-0-0/SC4_12-0-0_Codebook_de.pdf (accessed on 10 November 2022).
3.10	M	Adapted from FDZ-LIfBi (2021). Codebook NEPS Startkohorte 4 — Klasse 9. Schule und Ausbildung. Scientific Use File Version 12.0.0, pp. 3897, 3861. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC4/12-0-0/SC4_12-0-0_Codebook_de.pdf (accessed on 10 November 2022).
3.11	M	Adapted from FDZ-LIfBi (2021). Codebook NEPS Startkohorte 4 — Klasse 9. Schule und Ausbildung. Scientific Use File Version 12.0.0, pp. 6723-6724.

		https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC4/12-0-0/SC4_12-0-0_Codebook_de.pdf (accessed on 10 November 2022).
3.12	M	Item 3: Adapted from FDZ-LIfBi (2021). Codebook NEPS Startkohorte 4 — Klasse 9. Schule und Ausbildung. Scientific Use File Version 12.0.0, p. 160. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC4/12-0-0/SC4_12-0-0_Codebook_de.pdf (accessed on 10 November 2022).
3.13	M	Adapted from FDZ-LIfBi (2021). Codebook NEPS Startkohorte 4 — Klasse 9. Schule und Ausbildung. Scientific Use File Version 12.0.0, pp. 2941-2945. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC4/12-0-0/SC4_12-0-0_Codebook_de.pdf (accessed on 10 November 2022).
4.1	M	Beierlein, C., Kovaleva, A., Kemper, C. J., & Rammstedt, B. (2015). Kurzskala zur Erfassung der Risikobereitschaft (R-1). Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS). doi:10.6102/zis236.
4.2	M	European Social Survey (2016). ESS Round 8 Source Questionnaire. London: ESS ERIC Headquarters c/o City University London, p. 21. https://stessrelpubprodwe.blob.core.windows.net/data/round8/fieldwork/source/ESS8_source_questionnaires.pdf (accessed on 10 November 2022).
4.3	M	Abele, A. E., Stief, M. & Krüsken, J. (2002). Persönliche Ziele von Mathematikern beim Berufseinstieg. Ein Vergleich offener und geschlossener Erhebungsmethoden. Zeitschrift für pädagogische Psychologie, 16(3/4), pp. 193-205. https://doi.org/10.1024//1010-0652.16.34.193 (S. 196-197) (accessed on 16 December 2022).
4.4	O	GESIS – Leibniz-Institut für Sozialwissenschaften (o. J.). German Longitudinal Election Study. GLES 2013. Langfrist-Panel 2009-2013-2017. ZA5322, Version 1.1.0, Fragebogendokumentation, p. 22. https://access.gesis.org/dbk/60396 (accessed on 10 November 2022).
4.5	M	European Social Survey (2016). ESS Round 8 Source Questionnaire. London: ESS ERIC Headquarters c/o City University London, p. 10. https://stessrelpubprodwe.blob.core.windows.net/data/round8/fieldwork/source/ESS8_source_questionnaires.pdf (accessed on 10 November 2022).
4.6	M	Schupp, J. & Gerlitz, J.-Y. (2008). Big Five Inventory-SOEP (BFI-S). Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS). doi:10.6102/zis54.
4.7	M	Breen, R., van de Werfhorst, H. G. & Jæger, M. M. (2014). Deciding under Doubt: A Theory of Risk Aversion, Time Discounting Preferences, and Educational Decision-making. European Sociological Review, 30(2), 258–270. http://www.jstor.org/stable/24479880 (pp. 263-264) (accessed on 10 November 2022).
5.1	M	FDZ-LIfBi (2021). Codebook NEPS Startkohorte 4 — Klasse 9. Schule und Ausbildung. Scientific Use File Version 12.0.0, p. 4161. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC4/12-0-0/SC4_12-0-0_Codebook_de.pdf (accessed on 10 November 2022).
5.9	M	Beierlein, C., Kovaleva, A., László, Z., Kemper, C. J. & Rammstedt, B. (2015). Kurzskala zur Erfassung der Allgemeinen Lebenszufriedenheit (L-1). Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS). doi:10.6102/zis229.

Table7: Documentation of the origin of reused questions in Wave 1 (headteacher survey)

Question number	Original (O)/ Modified (M)	Citation
1.4	M	Mang, J., Ustjanzew, N., Leßke, I., Schiepe-Tiska, A. & Reiss, K. (2019). PISA 2015 Skalenhandbuch. Dokumentation der Erhebungsinstrumente. Münster: Waxmann. pp. 266-268. https://www.pisa.tum.de/fileadmin/w00bgi/www/Berichtsbaende_und_Zusammenfassungungen/Skalenhandbuch_PISA2015_openaccess.pdf (accessed on 10 November 2022).
1.6	M	Mang, J., Ustjanzew, N., Leßke, I., Schiepe-Tiska, A. & Reiss, K. (2019). PISA 2015 Skalenhandbuch. Dokumentation der Erhebungsinstrumente. Münster: Waxmann. p. 235. https://www.pisa.tum.de/fileadmin/w00bgi/www/Berichtsbaende_und_Zusammenfassungungen/Skalenhandbuch_PISA2015_openaccess.pdf (accessed on 10 November 2022).
1.8	M	Mang, J., Ustjanzew, N., Leßke, I., Schiepe-Tiska, A. & Reiss, K. (2019). PISA 2015 Skalenhandbuch. Dokumentation der Erhebungsinstrumente. Münster: Waxmann. p. 284. https://www.pisa.tum.de/fileadmin/w00bgi/www/Berichtsbaende_und_Zusammenfassungungen/Skalenhandbuch_PISA2015_openaccess.pdf (accessed on 10 November 2022).
2.1	M	Ditton, H. & Merz, D. (2013). Räumlichkeiten (Verfügbarkeit) - Schulleitung [Fragebogenskala: Version 1.0]. In: QualitätsSicherung in Schule und Unterricht - Fragebogenerhebung Erhebungszeitpunkt 1 (QuaSSU) [Skalenkollektion: Version 1.0]. Datenerhebung 2000. Frankfurt am Main: Forschungsdatenzentrum Bildung am DIPF. doi:10.7477/18:30:1. (see also https://www.fdz-bildung.de/skala.php?skala_id=4509&erhebung_id=30 (accessed on 10 November 2022)).
2.2	M	Mang, J., Ustjanzew, N., Leßke, I., Schiepe-Tiska, A. & Reiss, K. (2019). PISA 2015 Skalenhandbuch. Dokumentation der Erhebungsinstrumente. Münster: Waxmann. p. 237. https://www.pisa.tum.de/fileadmin/w00bgi/www/Berichtsbaende_und_Zusammenfassungungen/Skalenhandbuch_PISA2015_openaccess.pdf (accessed on 10 November 2022).
2.3	M	Mang, J., Ustjanzew, N., Leßke, I., Schiepe-Tiska, A. & Reiss, K. (2019). PISA 2015 Skalenhandbuch. Dokumentation der Erhebungsinstrumente. Münster: Waxmann. pp. 263-265. https://www.pisa.tum.de/fileadmin/w00bgi/www/Berichtsbaende_und_Zusammenfassungungen/Skalenhandbuch_PISA2015_openaccess.pdf (accessed on 10 November 2022).
2.4	M	Mang, J., Ustjanzew, N., Leßke, I., Schiepe-Tiska, A. & Reiss, K. (2019). PISA 2015 Skalenhandbuch. Dokumentation der Erhebungsinstrumente. Münster: Waxmann. p. 262. https://www.pisa.tum.de/fileadmin/w00bgi/www/Berichtsbaende_und_Zusammenfassungungen/Skalenhandbuch_PISA2015_openaccess.pdf (accessed on 10 November 2022).
2.5	M	Landkreis Rostock (Ed.) (2012). Anschlussperspektiven im Landkreis Rostock. Eine aktuelle Situationsbeschreibung am Übergang Schule-Beruf für den Zeitraum 2011/2012. PIW GmbH (2012). Bericht zum Auftrag des Projektes A-HA Anschluss Halten. „Durchführung einer Schulbefragung im Landkreis Rostock". p. 24. http://www.piw.de/Piw/Dokumente/Bericht%20Schulbefragung%2021-3-2012.pdf (accessed on 16 December 2022)
3.1	M	Landkreis Rostock (Ed.) (2012). Anschlussperspektiven im Landkreis Rostock. Eine aktuelle Situationsbeschreibung am Übergang Schule-Beruf für den Zeitraum 2011/2012. PIW GmbH (2012). Bericht zum Auftrag des Projektes A-HA Anschluss Halten. „Durchführung einer Schulbefragung im Landkreis Rostock". p. 24. http://www.piw.de/Piw/Dokumente/Bericht%20Schulbefragung%2021-3-2012.pdf (accessed on 16 December 2022)
3.4	M	Landkreis Rostock (Ed.) (2012). Anschlussperspektiven im Landkreis Rostock. Eine aktuelle Situationsbeschreibung am Übergang Schule-Beruf für den Zeitraum 2011/2012. PIW GmbH (2012). Bericht zum Auftrag des Projektes A-HA Anschluss Halten. „Durchführung einer Schulbefragung im Landkreis Rostock". p. 24. http://www.piw.de/Piw/Dokumente/Bericht%20Schulbefragung%2021-3-2012.pdf (accessed on 16 December 2022)

4.1	M	Fragebogen des PISA-Feldtests 2015 Deutschland, pp. 96-97 (unpublished).[48]
4.2	M	FDZ-LIfBi (2021). Codebook NEPS Startkohorte 4 — Klasse 9. Schule und Ausbildung. Scientific Use File Version 12.0.0, pp. 1395-1398. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC4/12-0-0/SC4_12-0-0_Codebook_de.pdf (accessed on 10 November 2022).

Table8: Documentation of the origin of reused questions in Wave 2

Question number	Original (O)/ Modified (M)	Citation
2.4	M	ManKobE (2014). Fragebogen für Auszubildende (Bayern). Mathematisch-naturwissenschaftliche Kompentenzen in der beruflichen Erstausbildung. Unveröffentlichte Skalendokumentation, p. 20.
2.6	M	ManKobE (2014). Fragebogen für Auszubildende (Bayern). Mathematisch-naturwissenschaftliche Kompentenzen in der beruflichen Erstausbildung. Unveröffentlichte Skalendokumentation, p. 20.
3.8	M	TOSCA 10. Welle 1 (Unveröffentlichte Skalendokumentation), p. 138.
3.9	M	FDZ-LIfBi (2023). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 18.0.0, pp. 2117-2130. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/18-0-0/SC5_18-0-0_Codebook_de.pdf (accessed on 16 August 2024).
3.10	M	Adapted from ManKobE (2014). Fragebogen für Auszubildende (Bayern). Mathematisch-naturwissenschaftliche Kompentenzen in der beruflichen Erstausbildung. Unveröffentlichte Skalendokumentation, p. 38; based on: Prenzel, M., Kristen, A., Dengler, P., Ettle, R., & Beer, T. (1996). Selbstbestimmt motiviertes und interessiertes Lernen in der kaufmännischen Erstausbildung. Zeitschrift für Berufs- und Wirtschaftspädagogik, Beiheft 13, 108-127.
3.11	M	FDZ-LIfBi (2023). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 18.0.0, pp. 2114-2122. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/18-0-0/SC5_18-0-0_Codebook_de.pdf (accessed on 16 August 2024).
3.12	M	Westermann, R., Heise, E. & Spies, K. (2018). FB-SZ-K. Kurzfragebogen zur Erfassung der Studienzufriedenheit [Verfahrensdokumentation, Fragebogen und Erläuterungen zum Fragebogen]. In Leibniz-Institut für Psychologie (ZPID) (Ed.), Open Test Archive. Trier: ZPID. https://doi.org/10.23668/psycharchives.4654
3.26	M	Items 6, 8, 9, 10, 12, 17, 18 and 20 are partly original and partly modified from TOSCA (Unveröffentlichte Skalendokumentation). The questionnaire can be requested from the IPN in Kiel.
4.5	M	Moschner B. & Gruber H. (2017). Erfassung epistemischer Überzeugung mit dem FEE, In: Bernholt A., Gruber H., & Moschner B. (Eds.), Wissen und Lernen, 1. Edition, Münster: Waxmann Verlag GmbH. pp. 17-37.
5.1	O	FDZ-LIfBi (2023). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 18.0.0, p. 2165. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/18-0-0/SC5_18-0-0_Codebook_de.pdf (accessed on 16 August 2024).
5.2	M	DIW Berlin / SOEP (Ed.) (2019) : SOEP-Core - 2016: Person (M1/M2 Wiederbefragte, CAPI, mit Verweis auf Variablen), SOEP Survey Papers, No. 634, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin, p. 35. https://www.econstor.eu/bitstream/10419/195345/1/1663422281.pdf (accessed on 28 October 2025) Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (2019). 21. Sozialerhebung. Fragebogen-Screenshots zu der Erhebung der wirtschaftlichen und

[48] A similar instrument was also used in: Brown, M., Altrichter, H., Nayir, F., Nortvedt, G., Burns, D., Fellner, M., Gloppen, S. K., Helm, C., McNamara, G., O'Hara, J., Punzenberger, B., Skedsmo, G. & Wiese, E. F. (2019). Classroom Assessment that Recognises Cultural Difference – A European Perspective. Final Report, Dublin (M. Felner, Trans.). DOI: 10.13140/RG.2.2.14346.52164/2, pp. 10–11 (Question number seven in the online survey questionnaire).

		sozialen Lage der Studierenden 2016: Deutsche und Bildungsinländer(innen). FDZ-DZHW, p. 81. https://metadata.fdz.dzhw.eu/public/files/instruments/inssy21-ins1$-2.0.0/attachments/ssy21_bi_Questionnaire_de.pdf (accessed on 26 August 2024).
5.4	M	FDZ-LIfBi (2023). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 18.0.0, p. 2163. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/18-0-0/SC5_18-0-0_Codebook_de.pdf (accessed on 16 August 2024).
6.1	M	Kantar Public (2018). SOEP-Core – 2017: Personenfragebogen, Stichproben A-L3. SOEP Survey Papers 563: Series A. Berlin: DIW/SOEP, p. 8. https://www.diw.de/documents/publikationen/73/diw_01.c.601702.de/diw_ssp0563.pdf (accessed on 26 August 2024).
6.2	M	Centers for Disease Control and Prevention (CDC). Behavioral Risk Factor Surveillance System Survey Questionnaire. Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention,1993, p. 3. https://www.cdc.gov/brfss/questionnaires/pdf-ques/93brfss.pdf (accessed on 26 August 2024). FDZ-LIfBi (2023). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 18.0.0, p. 1353. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/18-0-0/SC5_18-0-0_Codebook_de.pdf (accessed on 16 August 2024).
6.3	M	Klein, E. M., Brähler, E., Dreier, M., Reinecke, L., Müller, K. W., Schmutzer, G., Wölfing, K. & Beutel, M. E. (2016). The German version of the Perceived Stress Scale–psychometric characteristics in a representative German community sample. BMC psychiatry, 16(1), 159.
6.4	M	Moriarty, D.G., Zack, M.M., & Kobau, R. (2003). The Centers for Disease Control and Prevention's Healthy Days Measures – Population tracking of perceived physical and mental health over time. Health and Quality of Life Outcomes, 1, pp. 37-44. https://doi.org/10.1186/1477-7525-1-37 .
6.5	M	Moriarty, D.G., Zack, M.M., & Kobau, R. (2003). The Centers for Disease Control and Prevention's Healthy Days Measures – Population tracking of perceived physical and mental health over time. Health and Quality of Life Outcomes, 1, pp. 37-44. https://doi.org/10.1186/1477-7525-1-37 .
6.6	M	Moriarty, D.G., Zack, M.M., & Kobau, R. (2003). The Centers for Disease Control and Prevention's Healthy Days Measures – Population tracking of perceived physical and mental health over time. Health and Quality of Life Outcomes, 1, pp. 37-44. https://doi.org/10.1186/1477-7525-1-37 .
6.7	O	FDZ-LIfBi (2023). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 18.0.0, p. 1354. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/18-0-0/SC5_18-0-0_Codebook_de.pdf (accessed on 16 August 2024).
6.8	M	FDZ-LIfBi (2023). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 18.0.0, p. 1355. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/18-0-0/SC5_18-0-0_Codebook_de.pdf (accessed on 16 August 2024).
6.9	M	FDZ-LIfBi (2023). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 18.0.0, p. 1360. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/18-0-0/SC5_18-0-0_Codebook_de.pdf (accessed on 16 August 2024).
7.2	M	Demographische Standards. (2016). Eine gemeinsame Empfehlung des ADM Arbeitskreises Deutscher Markt- und Sozialforschungsinstitute e.V., der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V. (ASI) und des Statistischen Bundesamtes. Ausgabe 2016. 6. Überarbeitete und erweiterte Auflage, hrsg. vom Statistischen Bundesamt, Reihe: Statistik und Wissenschaft, Band 17. https://www.statistischebibliothek.de/mir/servlets/MCRFileNodeServlet/DEMonografie_derivate_00001549/Band17_DemographischeStandards1030817169004.pdf (accessed on 26 August 2024)
7.3	M	Statistische Ämter des Bundes und der Länder (2020). Mikrozensus 2011. Fragebogen Haushaltsbefragung. p. 8. https://www.zensus2011.de/SharedDocs/Downloads/DE/Fragebogen/Fragebogen_Haushaltebefragung.pdf?__blob=publicationFile&v=2 (accessed on 26 August 2024)

7.10	M	European Social Survey (ESS) (2012). ESS6 Main and Supplementary Questionnaire DE, p. 15. https://stessrelpubprodwe.blob.core.windows.net/data/round6/fieldwork/germany/ESS6_main_and_supplementary_questionnaire_DE.pdf (accessed on 27 August 2024).
7.11	M	GESIS - Leibniz-Institut für Sozialwissenschaften (2014). ALLBUS Allgemeine Bevölkerungsumfrage der Sozialwissenschaften 2014. ALLBUS 2014 Fragebogendokumentation. Material zu den Datensätzen mit den Studiennummern ZA5240 und ZA5241. Forschungsdatenzentrum ALLBUS, p. 40. https://dbk.gesis.org/DBKSearch/download.asp?id=55956 (accessed on 06 June 2023).
7.12	M	Beierlein, C., Kovaleva, A., László, Z., Kemper, C. J. & Rammstedt, B. (2015). Kurzskala zur Erfassung der Allgemeinen Lebenszufriedenheit (L-1). Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS). doi:10.6102/zis229.
7A	M	FDZ-LIfBi (2023). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 18.0.0, p. 2221. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/18-0-0/SC5_18-0-0_Codebook_de.pdf (accessed on 16 August 2024).

Table9: Documentation of the origin of reused questions in Waves 3a and 3b

survey	Question page	Original (O)/ Modified (M)	Citation
3a 3b	beg_03 beg_02	M	FDZ-LIfBi (2024). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 19.0.0, Bamberg, Germany, Leibniz Institute for Educational Trajectories, National Educational Panel Study, p. 859. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/19-0-0/SC5_19-0-0_Codebook_de.pdf (accessed on 30 March 2026).
3a 3b	beg_04 beg_03	M	FDZ-LIfBi (2024). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 19.0.0, Bamberg, Germany, Leibniz Institute for Educational Trajectories, National Educational Panel Study, p. 878. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/19-0-0/SC5_19-0-0_Codebook_de.pdf (accessed on 30 March 2026).
3a 3b	beg_08 beg_05	M	Danish Longitudinal Survey of Youth, vgl. u.a. Breen, R., van de Werfhorst, H. G., & Jæger, M. M. (2014). Deciding under doubt: A theory of risk aversion, time discounting preferences, and educational decisionmaking. European Sociological Review, 30(2), 258–270. Own translation.
3b	beg_06	O	Beierlein, C., Kovaleva, A., Kemper, C. J. & Rammstedt, B. (2015). Kurzskala zur Erfassung der Risikobereitschaft (R-1). Zusammenstellung sozialwissenschaftlicher Items und Skalen. https://doi.org/10.6102/zis236 .
3b	cor_01	O	FDZ-LIfBi (2020). Startkohorten 2-6. NEPS Corona und Bildung. Zusatzerhebung Mai 2020, Programmiervorlage CAWI. https://www.neps-data.de/Datenzentrum/Daten-und-Dokumentation/NEPS-C (accessed on 08 October 2026).
3b	cor_06	M	FDZ-LIfBi (2020). Startkohorten 2-6. NEPS Corona und Bildung. Zusatzerhebung Mai 2020, Programmiervorlage CAWI. https://www.neps-data.de/Datenzentrum/Daten-und-Dokumentation/NEPS-C (accessed on 08 October 2026).
3a 3b	stu_04 stu_03	M	Adapted from FDZ-LIfBi (2021). Codebook NEPS Startkohorte 4 — Klasse 9. Schule und Ausbildung. Scientific Use File Version 12.0.0, pp. 3897, 3861. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC4/12-0-0/SC4_12-0-0_Codebook_de.pdf (accessed on 10 November 2022).
3a 3b	stu_04 stu_04	M	Adapted from FDZ-LIfBi (2021). Codebook NEPS Startkohorte 4 — Klasse 9. Schule und Ausbildung. Scientific Use File Version 12.0.0,

			pp. 3897, 3861. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC4/12-0-0/SC4_12-0-0_Codebook_de.pdf (accessed on 10 November 2022).
3a 3b	stu_05 stu_05	M	Adapted from FDZ-LIfBi (2021). Codebook NEPS Startkohorte 4 — Klasse 9. Schule und Ausbildung. Scientific Use File Version 12.0.0, pp. 6723-6724. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC4/12-0-0/SC4_12-0-0_Codebook_de.pdf (accessed on 10 November 2022).
3a 3b	stu_05 stu_06	M	Adapted from FDZ-LIfBi (2021). Codebook NEPS Startkohorte 4 — Klasse 9. Schule und Ausbildung. Scientific Use File Version 12.0.0, pp. 6723-6724. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC4/12-0-0/SC4_12-0-0_Codebook_de.pdf (accessed on 10 November 2022).
3b	stu_07	O	Dickhäuser, O., Schöne, C., Spinath, B., Stiensmeier-Pelster, J. (2002). Die Skalen zum akademischen Selbstkonzept: Konstruktion und Überprüfung eines neuen Instrumentes. Zeitschrift für differentielle und diagnostische Psychologie 23 (4), pp. 393 - 405. https://doi.org/10.1024//0170-1789.23.4.393
3a 3b	stu_10 stu_08	M	FDZ-LIfBi (2024). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 19.0.0, Bamberg, Germany, Leibniz Institute for Educational Trajectories, National Educational Panel Study, pp. 2046, 2048-2049, 2052-2057, 2681, 2683. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/19-0-0/SC5_19-0-0_Codebook_de.pdf (accessed on 16 August 2024).
3a 3b	stu_11 stu_09	M	FDZ-LIfBi (2024). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 19.0.0, Bamberg, Germany, Leibniz Institute for Educational Trajectories, National Educational Panel Study, pp. 2059, 2061-2062, 2065-2066, 2068-2071, 2685. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/19-0-0/SC5_19-0-0_Codebook_de.pdf (accessed on 16 August 2024).
3b	stu_10	M	FDZ-LIfBi (2024). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 19.0.0, Bamberg, Germany, Leibniz Institute for Educational Trajectories, National Educational Panel Study, pp. 704-705. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/19-0-0/SC5_19-0-0_Codebook_de.pdf (accessed on 16 August 2024).
3b	stu_12	O	FDZ-LIfBi (2024). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 19.0.0, Bamberg, Germany, Leibniz Institute for Educational Trajectories, National Educational Panel Study, pp.1934, 1937, 1942-1945. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/19-0-0/SC5_19-0-0_Codebook_de.pdf (accessed on 16 August 2024).
3a 3b	stu_22 stu_15	O	FDZ-LIfBi (2024). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 19.0.0, Bamberg, Germany, Leibniz Institute for Educational Trajectories, National Educational Panel Study, pp. 2636-2640. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/19-0-0/SC5_19-0-0_Codebook_de.pdf (accessed on 16 August 2024).
3a 3b	stu_23 stu_16	O	FDZ-LIfBi (2024). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 19.0.0, Bamberg, Germany, Leibniz Institute for Educational Trajectories, National Educational Panel Study, pp. 2643-2645. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/19-0-0/SC5_19-0-0_Codebook_de.pdf (accessed on 16 August 2024).
3a 3b	stu_24 stu_17	M	FDZ-LIfBi (2024). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 19.0.0, Bamberg, Germany, Leibniz Institute for Educational Trajectories, National Educational Panel Study, p. 2646.

			https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/19-0-0/SC5_19-0-0_Codebook_de.pdf (accessed on 16 August 2024).
3a 3b	stu_25 stu_18	M	FDZ-LIfBi (2024). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 19.0.0, Bamberg, Germany, Leibniz Institute for Educational Trajectories, National Educational Panel Study, pp. 2647-2649. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/19-0-0/SC5_19-0-0_Codebook_de.pdf (accessed on 16 August 2024).
3a 3b	stu_27 stu_19	M	FDZ-LIfBi (2024). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 19.0.0, Bamberg, Germany, Leibniz Institute for Educational Trajectories, National Educational Panel Study, pp. 2650-2654. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/19-0-0/SC5_19-0-0_Codebook_de.pdf (accessed on 16 August 2024).
3a 3b	stu_28 stu_20	O	FDZ-LIfBi (2024). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 19.0.0, Bamberg, Germany, Leibniz Institute for Educational Trajectories, National Educational Panel Study, pp. 2655-2657. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/19-0-0/SC5_19-0-0_Codebook_de.pdf (accessed on 16 August 2024).
3a 3b	stu_29 stu_21	O	FDZ-LIfBi (2024). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 19.0.0, Bamberg, Germany, Leibniz Institute for Educational Trajectories, National Educational Panel Study, pp. 2658-2664. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/19-0-0/SC5_19-0-0_Codebook_de.pdf (accessed on 16 August 2024).
3a	stu_41	M	FDZ-LIfBi (2024). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 19.0.0, Bamberg, Germany, Leibniz Institute for Educational Trajectories, National Educational Panel Study, pp. 2636-2640. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/19-0-0/SC5_19-0-0_Codebook_de.pdf (accessed on 16 August 2024).
3a	stu_44	M	Custom construct based on: Studierfähigkeit - institutionelle Förderung und studienrelevante Heterogenität (StuFHe). https://www.ew.uni-hamburg.de/forschung/forschungsprojekte/stufhe.html (accessed on 26 August 2024)
3a, 3b	mas_01	M	Custom construct based on: FDZ-LIfBi (2024). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 19.0.0, Bamberg, Germany, Leibniz Institute for Educational Trajectories, National Educational Panel Study, pp. 820-824. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/19-0-0/SC5_19-0-0_Codebook_de.pdf (accessed on 16 August 2024).
3a, 3b	gsh_01	O	Centers for Disease Control and Prevention (CDC). Behavioral Risk Factor Surveillance System Survey Questionnaire. Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention,1993, Page 3. https://www.cdc.gov/brfss/questionnaires/pdf-ques/93brfss.pdf (accessed on 26 August 2024).
3b	gsh_02	O	Klein, E. M., Brähler, E., Dreier, M., Reinecke, L., Müller, K. W., Schmutzer, G., Wölfing, K. & Beutel, M. E. (2016). The German version of the Perceived Stress Scale–psychometric characteristics in a representative German community sample. BMC psychiatry, 16(1), 159.
3b	gsh_03	M	Moriarty, D.G., Zack, M.M., & Kobau, R. (2003). The Centers for Disease Control and Prevention's Healthy Days Measures –

			Population tracking of perceived physical and mental health over time. Health and Quality of Life Outcomes, 1, pp. 37-44. https://doi.org/10.1186/1477-7525-1-37
3b	gsh_04	M	Moriarty, D.G., Zack, M.M., & Kobau, R. (2003). The Centers for Disease Control and Prevention's Healthy Days Measures – Population tracking of perceived physical and mental health over time. Health and Quality of Life Outcomes, 1, pp. 37-44. https://doi.org/10.1186/1477-7525-1-37
3b	gsh_05	M	Moriarty, D.G., Zack, M.M., & Kobau, R. (2003). The Centers for Disease Control and Prevention's Healthy Days Measures – Population tracking of perceived physical and mental health over time. Health and Quality of Life Outcomes, 1, pp. 37-44. https://doi.org/10.1186/1477-7525-1-37
3a	msg_01	O	FDZ-LIfBi (2024). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 19.0.0, Bamberg, Germany, Leibniz Institute for Educational Trajectories, National Educational Panel Study, p. 1290. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/19-0-0/SC5_19_0-0_Codebook_de.pdf (accessed on 03 March 2026).
3b	msg_01	O	Beierlein, C., Kovaleva, A., Kemper, C. J. & Rammstedt, B. (2014). Allgemeine Selbstwirksamkeit Kurzskala (ASKU). Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS). https://doi.org/10.6102/zis35
3b	msg_02	M	FDZ-LIfBi. (2025). Codebook NEPS Startkohorte 4 — Klasse 9 Schule und Ausbildung — Scientific Use File Version 19.0.0, Bamberg, Germany, Leibniz Institute for Educational Trajectories, National Educational Panel Study, p. 4641.https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/19-0-0/SC5_19_0-0_Codebook_de.pdf (accessed on 03 March 2026). Adapted from: Lundmark, S., Gilljam, M. & Dahlberg, S. (2016). Measuring Generalized Trust. An Examination of Question Wording and the Number of Scale Points. Public Opinion Quarterly, 80(1), 26-43. https://doi.org/10.1093/poq/nfv042
3b	msg_03	M	Lundmark, S., Gilljam, M. & Dahlberg, S. (2016). Measuring Generalized Trust. An Examination of Question Wording and the Number of Scale Points. Public Opinion Quarterly, 80(1), 26-43. https://doi.org/10.1093/poq/nfv042
3b	msg_05	M	Moschner B. & Gruber H. (2017). Erfassung epistemischer Überzeugung mit dem FEE, In: Bernholt A., Gruber H., & Moschner B. (Eds.), Wissen und Lernen, 1. Edition, Münster: Waxmann Verlag GmbH. pp. 17-37.
3a 3b	msg_10 msg_20	M	Beierlein, C., Kovaleva, A., László, Z., Kemper, C. J., & Rammstedt, B. (2015). Kurzskala zur Erfassung der Allgemeinen Lebenszufriedenheit (L-1). Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS). https://doi.org/10.6102/zis229
Episode module	stXX19	M	Adapted from: Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (2017). 20. Sozialerhebung (2012). Fragebogen. Version 1.0.0. FDZ-DZHW, p. 14 (Question 49). https://metadata.fdz.dzhw.eu/public/files/instruments/ins-ssy20-ins1$/attachments/ssy20_Questionnaire_de.pdf (accessed on 09 October 2026)
Episodic module	emXX06	M	Statistisches Bundesamt. (2011). Zensus 2011: Haushaltfragebogen. Questions 45 and 46. https://unstats.un.org/unsd/demographic/sources/census/quest/DEU2011deHh.pdf (accessed on 16 March 2026)
Episode module	emXX12	M	Demographische Standards. (2016). Eine gemeinsame Empfehlung des ADM Arbeitskreises Deutscher Markt- und Sozialforschungsinstitute e.V., der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V. (ASI) und des Statistischen Bundesamtes. Ausgabe 2016, pp. 46f. 6. überarbeitete und erweiterte Auflage, hrsg. vom Statistischen Bundesamt, Reihe: Statistik und Wissenschaft, Band 17.

			https://www.statistischebibliothek.de/mir/servlets/MCRFileNodeServlet/DEMonografie_derivate_00001549/Band17_DemographischeStandards1030817169004.pdf (accessed on 30 March 2026)
Episode module	emXX40	M	Adapted from: Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (2022). DZHW-Absolventenpanel 2009. Fragebogen zu der Erhebung der Absolvent(inn)enkohorte 2009 (3. Befragungswelle). Version 2.0.0. FDZ-DZHW, p. 34 (p314). https://metadata.fdz.dzhw.eu/public/files/instruments/insgra2009-ins6$/attachments/gra2009_W3_Main_Questionnaire_de.pdf (accessed on 08 October 2026).
Episode module	emXX46	M	Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (2021). DZHW-Absolventenpanel 2013. Fragebogen zu der Erhebung der Absolvent(inn)enkohorte 2013 (2. Befragungswelle). Version 3.0.0. FDZ-DZHW, p. 31 (p316). https://metadata.fdz.dzhw.eu/public/files/instruments/insgra2013-ins2$-3.0.0/attachments/gra2013_W2_Questionnaire_de.pdf (accessed on 30 March 2026)
Episode module	emXX47	M	Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (2021). DZHW-Absolventenpanel 2013. Fragebogen zu der Erhebung der Absolvent(inn)enkohorte 2013 (2. Befragungswelle). Version 3.0.0. FDZ-DZHW, p. 32 (p317). https://metadata.fdz.dzhw.eu/public/files/instruments/insgra2013-ins2$-3.0.0/attachments/gra2013_W2_Questionnaire_de.pdf (accessed on 30 March 2026)
Episode module	seXX06	M	Demographische Standards. (2016). Eine gemeinsame Empfehlung des ADM Arbeitskreises Deutscher Markt- und Sozialforschungsinstitute e.V., der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V. (ASI) und des Statistischen Bundesamtes. Ausgabe 2016, pp. 46f. 6. überarbeitete und erweiterte Auflage, hrsg. vom Statistischen Bundesamt, Reihe: Statistik und Wissenschaft, Band 17. https://www.statistischebibliothek.de/mir/servlets/MCRFileNodeServlet/DEMonografie_derivate_00001549/Band17_DemographischeStandards1030817169004.pdf (accessed on 16 March 2026)
Episode module	seXX11	M	Demographische Standards. (2016). Eine gemeinsame Empfehlung des ADM Arbeitskreises Deutscher Markt- und Sozialforschungsinstitute e.V., der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V. (ASI) und des Statistischen Bundesamtes. Ausgabe 2016, pp. 46f. 6. überarbeitete und erweiterte Auflage, hrsg. vom Statistischen Bundesamt, Reihe: Statistik und Wissenschaft, Band 17. https://www.statistischebibliothek.de/mir/servlets/MCRFileNodeServlet/DEMonografie_derivate_00001549/Band17_DemographischeStandards1030817169004.pdf (accessed on 16 March 2026)
Episode module	seXX12	M	Demographische Standards. (2016). Eine gemeinsame Empfehlung des ADM Arbeitskreises Deutscher Markt- und Sozialforschungsinstitute e.V., der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V. (ASI) und des Statistischen Bundesamtes. Ausgabe 2016, pp. 46f. 6. überarbeitete und erweiterte Auflage, hrsg. vom Statistischen Bundesamt, Reihe: Statistik und Wissenschaft, Band 17. https://www.statistischebibliothek.de/mir/servlets/MCRFileNodeServlet/DEMonografie_derivate_00001549/Band17_DemographischeStandards1030817169004.pdf (accessed on 16 March 2026)
Episode module	seXX15	M	Statistisches Bundesamt. (2011). Zensus 2011: Haushaltfragebogen. Questions 45 and 46. https://unstats.un.org/unsd/demographic/sources/census/quest/DEU2011deHh.pdf (accessed on 16 March 2026)
Episode module	iXX15	M	Adapted from: Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (2022). DZHW-Absolventenpanel 2009. Fragebogen zu der Erhebung der Absolvent(inn)enkohorte 2009 (3. Befragungswelle). Version 2.0.0. FDZ-DZHW, p. 34 (p314). https://metadata.fdz.dzhw.eu/public/files/instruments/ins-

			gra2009-ins6$/attachments/gra2009_W3_Main_Questionnaire_de.pdf (accessed on 08 October 2026).
Episode module	iXX18	M	FDZ-LIfBi (2024). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 19.0.0, Bamberg, Germany, Leibniz Institute for Educational Trajectories, National Educational Panel Study, pp. 4117, 4126. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/19-0-0/SC5_19_0-0_Codebook_de.pdf (accessed on 03 March2026).
Episode module	iXX20	M	FDZ-LIfBi (2024). Codebook. NEPS Startkohorte 5 - Studierende. Hochschulstudium und Übergang in den Beruf. Scientific Use File Version 19.0.0, Bamberg, Germany, Leibniz Institute for Educational Trajectories, National Educational Panel Study, p. 4112. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/19-0-0/SC5_19_0-0_Codebook_de.pdf (accessed on 03 March 2026).
Episode module	tXX05	M	Adapted from: Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (2022). DZHW-Absolventenpanel 2009. Fragebogen zu der Erhebung der Absolvent(inn)enkohorte 2009 (2. Befragungswelle). Version 2.0.0. FDZ-DZHW, p. 12 (6.2). https://metadata.fdz.dzhw.eu/public/files/instruments/ins-gra2009-ins2$/attachments/gra2009_W2_Main_PAPI_Questionnaire_de.pdf (accessed on 9 October 2026) FDZ-LIfBi. (2026). Codebook NEPS Startkohorte 6 — Erwachsene Bildung im Erwachsenenalter und lebenslanges Lernen Scientific Use File Version 19.0.0, Bamberg, Germany, Leibniz Institute for Educational Trajectories, National Educational Panel Study, p. 3636. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC6/17-0-0/NEPS_SC6_Codebook_17-0-0_de.pdf (accessed on 9 October 2026)