

NEPS

National Educational Panel Study

Ariane Würbach

Samples, Weights and Nonresponse

NEPS Starting Cohort 3 — Grade 5

*Paths Through Lower Secondary School —
Educational Pathways of Students in Grade 5
and Higher*

Wave 12



LifBi



LEIBNIZ INSTITUTE FOR
EDUCATIONAL TRAJECTORIES

Copyrighted Material
Leibniz Institute for Educational Trajectories (LifBi)
Wilhelmsplatz 3, 96047 Bamberg
Director: Prof. Dr. Cordula Artelt
Administrative Director: Dr. Stefan Echinger
Bamberg; November 24, 2022

Samples, Weights, and Nonresponse: the Sample of Starting Cohort 3 of the National Educational Panel Study (Wave 12)

Ariane Würbach

Leibniz Institute for Educational Trajectories

Technical Report referring to [DOI:10.5157/NEPS:SC3:12.0.0](https://doi.org/10.5157/NEPS:SC3:12.0.0)

E-mail address:

statistik@lifbi.de

Samples, Weights, and Nonresponse: the Sample of Starting Cohort 3 of the National Educational Panel Study (Wave 12)

1 Prequel

This report complements NEPS Survey Paper No. 63 (Steinhauer & Zinn, 2016a) and gives details on Wave 12 of Starting Cohort 3 (SC3) of the National Educational Panel Study (NEPS). It refers to the Scientific Use File (SUF; [DOI:10.5157/NEPS:SC3:12.0.0](https://doi.org/10.5157/NEPS:SC3:12.0.0)). SC3 focuses on students in Grade 5 and their pathway through lower secondary education. The original sample consists of a main sample of Grade 5 students in regular schools and special-needs schools with a supplement covering students with a migration background from Turkey and the former Soviet Union. Due to the Federal-State-specific timing in transition in lower secondary education in regular schools a refreshment sample was drawn for students attending Grade 7.

To provide weights for the original samples as well as for the refreshment sample the different processes leading to the participation decision in a certain wave have to be considered. These decision processes include the schools initial decision to participate in the survey, the students initial decision to participate in the survey, and the students successive decisions to participate in each wave again. The schools initial decision to participate enters a nonresponse adjusted design weight on the institutional level. The students initial decision to participate enters a nonresponse adjusted design weight on the individual level. The successive decisions of a student to participate in a certain wave enter the corresponding wave-specific cross-sectional and longitudinal weights.

The students willing to participate in the panel study (i.e., the panel members) are followed up over time. In the progress of the panel it is possible that students cannot be surveyed within their institutional context for several reasons. For example, because they switch to another school, or because the school decides to refuse further cooperation. In these cases students are surveyed in an individual context, that is, the questionnaires are sent to their home address. Surveying students in this individual context is referred to as the field of individual retracking.

Table 1 illustrates the number of students according to the sample they originally belong to and their participation status by wave. The table gives details on the size of panel cohort over time. The column "Used sample" is split up into "Participants", "Temporary dropouts", and "Final dropouts (in Wave)", displaying the students status by the end of the wave. Finally, the last column presents the number of students withdrawing their panel consent between two rounds of survey waves.

For details on the sampling design and the derivation of design weights see Steinhauer et al. (2015). Details on calculating wave-specific nonresponse adjustments can be found in Steinhauer and Zinn (2016a) for Waves 1 to 3, in Steinhauer and Zinn (2016b) for Waves 1 to 5, in Steinhauer (2017) for Waves 6 and 7, in Steinhauer (2019) for Wave 8, in Hammon and Landrock (2019) for Wave 9, in Schnapp (2020) for Wave 10, and in Schnapp (2021) for Wave 11.

2 Changes compared to the previous version

Weights for Wave 12 (Study B134) have been appended.

Table 1: Panel progress of SC3 by wave.

Wave (Time)	Study number		Panel Cohort			Status at the end of the wave			
			Total size	Not used	Used sample	Participants	Temporary dropout	Final dropout (in wave)	Final dropout (after wave)
1 (2010/2011)	(A28, A56, A63)	Main	6112	0	6112	5778	334	0	13
2 (2011/2012)	(A29, A57)	Main	6099	0	6099	5538	560	1	8
3 (2012/2013)	(A30, A30A, A58)	All	8295	0	8295	7277	989	29	10
		Main	6090	0	6090	5131	930	29	10
		Refr.	2205	0	2205	2146	59	0	0
4 (2013/2014)	(A31, A59)	All	8256	0	8256	6718	1505	33	580
		Main	6051	0	6051	4783	1249	19	^a 580
		Refr.	2205	0	2205	1935	256	14	0
5 (2014/2015)	(A94)	All	7643	0	7643	5778	1625	240	0
		Main	5452	0	5452	4001	1273	178	0
		Refr.	2191	0	2191	1777	352	62	0
6 (Spring 2015)	(A98)	All	7403	0	7403	5586	1740	77	2
		Main	5274	0	5274	3920	1293	61	2
		Refr.	2129	0	2129	1666	447	16	0
7 (2015/2016)	(A99, B106)	All	7324	244	7080	5492	1543	45	29
		Main	5211	153	5058	3925	1104	29	21
		Refr.	2113	91	2022	1567	439	16	8
8 (2016/2017)	(A100, B107)	All	7250	65	7185	5263	1562	360	244
		Main	5161	42	5119	3767	1095	257	188
		Refr.	2089	23	2066	1496	467	103	56
9 (2017/2018)	(A101, B108)	All	6646	1	6645	4988	1184	473	513
		Main	4716	0	4716	3590	812	314	346
		Refr.	1930	1	1929	1398	372	159	167
10 (2018/2019)	(B132)	All	5660	0	5660	3846	1516	298	238
		Main	4056	0	4056	2774	1081	201	172
		Refr.	1604	0	1604	1072	435	97	6

Table 1: Panel progress of SC3 by wave.

Wave (Time)	Study number		Panel Cohort			Status at the end of the wave			
			Total size	Not used	Used sample	Participants	Temporary dropout	Final dropout (in wave)	Final dropout (after wave)
11 (2019/2020)	(B133)	All	5124	0	5124	3292	1629	203	799
		Main	3683	0	3683	2393	1161	129	567
		Refr.	1441	0	1441	899	468	74	232
12 (2020/2021)	(B134)	All	4122	0	4122	2924	1131	67	71
		Main	2987	0	2987	2138	805	44	567
		Refr.	1135	0	1135	786	326	23	232

^a special-need students are excluded from the panel cohort after Wave 4.

3 Participation in Wave 12

To account for the wave-specific participation decision of students response propensity re-weighting is used to provide corresponding weights. To model binary participation decisions a model with a probit link function is used and adapted via a stepwise selection for Wave 8 and subsequent waves, see Steinhauer and Zinn (2016a) and Steinhauer and Zinn (2016b) for origins and details. By Wave 12 the panel cohort has reduced to 4,122 students, see Table 1. Like in NEPS Starting Cohort 4 these students left their schools and thus are surveyed individually. The significant coefficients for the estimated models are displayed in Table 2.

Participation in previous waves significantly increases the propensity the participate in Wave 12, for both sample groups. Additionally, and again for both sample groups, students with native language other than German have a significantly decreased the propensity to participate. Students of the main sample exhibit a significantly increased propensity to participate when they are in the younger age group, compared to the older half.

Table 2: Models estimating the individual participation propensities for students in Wave 12 of SC3 used to derive adjustment factors for adjusted wave-specific cross-sectional and longitudinal weights.

	Wave 12 Main Sample	Refreshment Sample
(Intercept)	−1.811*** (0.399)	−2.234*** (0.235)
Age group: younger half	0.162* (0.085)	
Native language: other	−0.234* (0.125)	−0.291*** (0.091)
Native language: missing	1.150 (0.799)	0.271 (0.244)
Gender: female	−0.120 (0.084)	
Student participated in Wave 3	0.633* (0.349)	
Student participated in Wave 1		0.274* (0.151)
Student participated in Wave 2		0.209* (0.125)
Student participated in Wave 5	0.403*** (0.134)	0.186** (0.085)
Student participated in Wave 6		0.144* (0.084)
Student participated in Wave 7		0.221*** (0.084)
Student participated in Wave 8		0.153 (0.094)
Student participated in Wave 10	0.348** (0.146)	0.756*** (0.087)
Student participated in Wave 11	1.352*** (0.100)	1.446*** (0.064)
Observations	1,135	2,901

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$; standard errors are given in parentheses. To model individual participation, the `glm` function with a probit link provided in R (R Core Team, 2020) was used. AIC based backward selection was used and only significant coefficients are reported. Reference categories are: Age group (older half), Native language (German), Gender (male), Student participated in Wave t (no).

4 Summary of Weights

Various kinds of weights for students together with design information are provided by NEPS. Table 3 summarizes the design information and the different weights provided by SUF release version DOI:10.5157/NEPS:SC3:12.0.0. Besides individual/target (ID_t) and institutional (ID_i) identifiers, design information for the entire cohort is made available.¹ This information covers the study number corresponding to the first survey in which a student had been surveyed, the explicit sampling strata (stratum_exp) as well as the implicit sampling strata. Variables used for implicit stratification are "school type" (stratum_imp1), "federal state" (stratum_imp2_R), "regional classification" (stratum_imp3_R) and "funding" (stratum_imp4_R).

Nonresponse adjusted design weights on the institutional (w_i) and the individual (w_t) level are given for the entire cohort.² For all participants in a particular wave, cross-sectional weights are provided. With respect to panel progress longitudinal weights are also available. Data from Official Statistics (Statistisches Bundesamt, Fachserie 11, Reihe 1, 2010/11) regarding the gender ratio in different school types of different Federal States has been used for initial raking.

To ease statistical analysis, all weights apart from the pure design weights (Wave 1 and Wave 3) are provided in a trimmed and standardized form. Summary statistics for all kind of weights provided are given in Table 4.

Table 3: Variables included in the weighting data set for SC3 SUF version 12.0.0.

Variable	Applies to	Content
ID_t	8317	Identifier for target person
ID_i	8317	Identifier for the institution
<i>Design information</i>		
tstud_st	8317	Study number the target person was first surveyed in
sample	8317	Part of the sample the target person belongs to
stratum_exp	8317	Explicit stratum referring to school
stratum_imp1_R	8317	Implicit stratum (school type according to sampling frame)
stratum_imp2_R	8317	Implicit stratum (federal state according to sampling frame)
stratum_imp3_R	8317	Implicit stratum (regional classification according to sampling frame)
stratum_imp4_R	8317	Implicit stratum (funding according to sampling frame)
tx80113_R	7670	Total number of classes in grade 8 as reported by official statistics
tx80114_R	7670	Total number of students in grade 8 as reported by official statistics
<i>Design weights adjusted for initial nonresponse</i>		
w_i	8317	Design weight for institution
w_t	8317	Design weight for target
w_t_cal	5283	Design weight for target, calibrated
w_t3_cal	8054	Design weight for target in Wave 3, calibrated

¹Due to data protection, this information is not available in the download version of the SUF.

²The institutional weight as well as the explicit and implicit stratification variables belong to the institution and thus are equal for all cases within the institution.

Table 3: Variables included in the weighting data set for SC3 SUF version 12.0.0. (continued)

Variable	Applies to	Content
<i>Weights for targets adjusted for wave-specific nonresponse</i>		
w_t1	5559	Cross-sectional weight for targets participating in Wave 1
w_t2	5330	Cross-sectional weight for targets participating in Wave 2
w_t3	7111	Cross-sectional weight for targets participating in Wave 3
w_t4	6581	Cross-sectional weight for targets participating in Wave 4
w_t5	5648	Cross-sectional weight for targets participating in Wave 5
w_t6	5465	Cross-sectional weight for targets participating in Wave 6
w_t7	5367	Cross-sectional weight for targets participating in Wave 7
w_t8	5139	Cross-sectional weight for targets participating in Wave 8
w_t9	4870	Cross-sectional weight for targets participating in Wave 9
w_t10	3766	Cross-sectional weight for targets participating in Wave 10
w_t11	3231	Cross-sectional weight for targets participating in Wave 11
w_t12	2883	Cross-sectional weight for targets participating in Wave 12
w_t1to2	5070	Longitudinal weight for targets participating in Wave 1 to 2
w_t1to3	4514	Longitudinal weight for targets participating in Wave 1 to 3
w_t1to4	4027	Longitudinal weight for targets participating in Wave 1 to 4
w_t1to5	3203	Longitudinal weight for targets participating in Wave 1 to 5
w_t1to6	2919	Longitudinal weight for targets participating in Wave 1 to 6
w_t1to7	2604	Longitudinal weight for targets participating in Wave 1 to 7
w_t1to8	2228	Longitudinal weight for targets participating in Wave 1 to 8
w_t1to9	1947	Longitudinal weight for targets participating in Wave 1 to 9
w_t1to10	1483	Longitudinal weight for targets participating in Wave 1 to 10
w_t1to11	1185	Longitudinal weight for targets participating in Wave 1 to 11
w_t1to12	1019	Longitudinal weight for targets participating in Wave 1 to 12
w_t3to4	6288	Longitudinal weight for targets participating in Wave 3 to 4
w_t3to5	5119	Longitudinal weight for targets participating in Wave 3 to 5
w_t3to6	4601	Longitudinal weight for targets participating in Wave 3 to 6
w_t3to7	4027	Longitudinal weight for targets participating in Wave 3 to 7
w_t3to8	3361	Longitudinal weight for targets participating in Wave 3 to 8
w_t3to9	2889	Longitudinal weight for targets participating in Wave 3 to 9
w_t3to10	2185	Longitudinal weight for targets participating in Wave 3 to 10
w_t3to11	1740	Longitudinal weight for targets participating in Wave 3 to 11
w_t3to12	1475	Longitudinal weight for targets participating in Wave 3 to 12
<i>Weights for targets and parents adjusted for wave-specific nonresponse</i>		
w_tp1	3550	Cross-sectional weight for joint participation in Wave 1
w_tp2	3307	Cross-sectional weight for joint participation in Wave 2
w_tp3	4248	Cross-sectional weight for joint participation in Wave 3
w_tp4	2621	Cross-sectional weight for joint participation in Wave 4
w_tp6	2776	Cross-sectional weight for joint participation in Wave 6
w_tp1to2	3042	Longitudinal weight for joint participation in Wave 1 to 2
w_tp1to3	2544	Longitudinal weight for joint participation in Wave 1 to 3
w_tp1to4	2107	Longitudinal weight for joint participation in Wave 1 to 4
w_tp1to6	1511	Longitudinal weight for joint participation in Wave 1 to 6 (without 5)
w_tp3to4	3404	Longitudinal weight for joint participation in Wave 3 to 4

Table 3: Variables included in the weighting data set for SC3 SUF version 12.0.0. (continued)

Variable	Applies to	Content
w_tp3to6	2298	Longitudinal weight for joint participation in Wave 3 to 6 (without 5)
<i>Weights for target of the Corona-CAWI adjusted for nonresponse</i>		
w_tC	1021	Cross-sectional weight for targets participating in Corona-CAWI
w_tC_cal	1021	Calibrated cross-sectional weight for targets participating in Corona-CAWI
w_t1toC	398	Longitudinal weight for targets participating in Wave 1 to Corona-CAWI
w_t3toC	574	Longitudinal weight for targets participating in Wave 3 to Corona-CAWI

Table 4: Summary statistics for all weights provided.

Label of weight	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
w_i	0.9428	71.5588	94.8200	133.8641	126.4267	17545.3125	
w_t	0.9491	115.3820	163.3531	349.8744	244.9403	432548.3206	
w_t_cal	23.3228	80.2816	109.8667	143.6241	165.4758	14735.9678	3034
w_t3_cal	2.3708	47.0307	73.6537	99.4404	112.3628	17252.6377	
w_t1	0.0317	0.5634	0.8001	1.0000	1.2322	3.7276	2758
w_t2	0.0282	0.4919	0.6978	1.0000	1.1431	4.4564	2987
w_t3	0.0221	0.4030	0.6387	1.0000	1.0885	4.8685	1206
w_t4	0.0211	0.4163	0.6502	1.0000	1.0826	4.7776	1736
w_t5	0.0363	0.3117	0.4784	1.0000	0.8758	5.3877	2669
w_t6	0.0205	0.1835	0.2986	1.0000	0.6611	5.7085	2852
w_t7	0.0113	0.1103	0.1967	1.0000	0.8453	5.7251	2950
w_t8	0.0067	0.0725	0.1582	1.0000	1.0698	5.7109	3178
w_t9	0.0050	0.0593	0.1595	1.0000	1.0711	5.7028	3447
w_t10	0.0043	0.0520	0.1530	1.0000	1.0445	5.7200	4551
w_t11	0.0033	0.0430	0.1481	1.0000	0.9347	5.7678	5086
w_t12	0.0026	0.0347	0.1520	1.0000	0.8460	5.8123	5434
w_t1to2	0.0328	0.5596	0.7905	1.0000	1.2479	3.7075	3247
w_t1to3	0.0319	0.5465	0.7814	1.0000	1.2514	3.7326	3803
w_t1to4	0.0318	0.5104	0.7470	1.0000	1.2264	3.9322	4290
w_t1to5	0.1650	0.4917	0.7276	1.0000	1.2101	3.9450	5114
w_t1to6	0.1605	0.4780	0.7169	1.0000	1.2033	4.0141	5398
w_t1to7	0.1645	0.4605	0.6983	1.0000	1.1778	4.1748	5713
w_t1to8	0.1512	0.4362	0.6790	1.0000	1.1679	4.2910	6089
w_t1to9	0.1408	0.4199	0.6635	1.0000	1.1628	4.3769	6370
w_t1to10	0.1313	0.4032	0.6683	1.0000	1.1613	4.4307	6834
w_t1to11	0.1271	0.3987	0.6678	1.0000	1.1623	4.4158	7132
w_t1to12	0.1263	0.3947	0.6619	1.0000	1.1388	4.4775	7298
w_t3to4	0.0256	0.4916	0.7659	1.0000	1.2023	4.1130	2029
w_t3to5	0.0603	0.4916	0.7395	1.0000	1.1911	4.1558	3198
w_t3to6	0.0576	0.4716	0.7204	1.0000	1.1834	4.2559	3716

Table 4: Summary statistics for all weights provided. (continued)

Label of weight	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
w_t3to7	0.0546	0.4597	0.7022	1.0000	1.1699	4.3740	4290
w_t3to8	0.0501	0.4262	0.6636	1.0000	1.1562	4.4987	4956
w_t3to9	0.0475	0.4146	0.6552	1.0000	1.1346	4.5720	5428
w_t3to10	0.0464	0.4063	0.6479	1.0000	1.1521	4.6009	6132
w_t3to11	0.0461	0.4054	0.6411	1.0000	1.1521	4.6062	6577
w_t3to12	0.0465	0.4026	0.6364	1.0000	1.1443	4.6086	6842
w_tp1	0.1851	0.6301	0.8321	1.0000	1.2171	3.0507	4767
w_tp2	0.1286	0.4134	0.5708	1.0000	0.9465	5.2279	5010
w_tp3	0.0270	0.3205	0.4890	1.0000	0.8073	5.4755	4069
w_tp4	0.0308	0.0885	0.1415	1.0000	0.2976	5.9171	5696
w_tp6	0.0174	0.1162	0.2059	1.0000	0.5617	5.8297	5541
w_tp1to2	0.1961	0.6161	0.8228	1.0000	1.2401	3.0645	5275
w_tp1to3	0.2213	0.5872	0.8070	1.0000	1.2381	3.2123	5773
w_tp1to4	0.2118	0.5611	0.7835	1.0000	1.2549	3.4148	6210
w_tp1to6	0.1884	0.5175	0.7527	1.0000	1.2720	3.6709	6806
w_tp3to4	0.0253	0.4568	0.6955	1.0000	1.1361	4.4156	4913
w_tp3to6	0.0700	0.4243	0.6735	1.0000	1.1082	4.5781	6019
w_tC	0.0040	0.0440	0.1580	1.0000	0.9290	5.7790	7296
w_tC_cal	0.0010	0.0400	0.1430	1.0000	0.8680	21.9660	7296
w_t1toC	0.1450	0.3750	0.6550	1.0000	1.1840	4.5780	7296
w_t3toC	0.0490	0.3540	0.5790	1.0000	1.2090	4.8220	7296

For further information on weighting please contact statistik@lifbi.de.

Acknowledgements

This paper uses data from the National Educational Panel Study (NEPS; see Blossfeld and Roßbach, 2019). The NEPS is carried out by the Leibniz Institute for Educational Trajectories (LifBi, Germany) in cooperation with a nationwide network.

- Blossfeld, H.-P., & Roßbach, H. G. (Eds.). (2019). *Education as a lifelong process: The German National Educational Panel Study (NEPS). Edition ZfE (2nd)*. Springer VS.
- Hammon, A., & Landrock, U. (2019). *Samples, Weights, and Nonresponse: the Sample of Starting Cohort 3 of the National Educational Panel Study (Wave 9)*. Leibniz Institute for Educational Trajectories, National Educational Panel Study. Bamberg. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC3/9-0-0/SC3_9-0-0_W9.pdf
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org/>
- Schnapp, T. (2020). *Samples, Weights, and Nonresponse: the Sample of Starting Cohort 3 of the National Educational Panel Study (Wave 10)*. Leibniz Institute for Educational Trajectories, National Educational Panel Study. Bamberg. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC3/10-0-0/SC3_10-0-0_W10.pdf
- Schnapp, T. (2021). *Samples, Weights, and Nonresponse: the Sample of Starting Cohort 3 of the National Educational Panel Study (Wave 11)*. Leibniz Institute for Educational Trajectories, National Educational Panel Study. Bamberg. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC3/11-0-0/SC3_11-0-0_W11.pdf
- Statistisches Bundesamt. (2011). *Bildung und Kultur - Allgemeinbildende Schulen: Schuljahr 2010/2011* (Fachserie 11 Reihe 1). Statistisches Bundesamt. Wiesbaden. https://www.statistischebibliothek.de/mir/servlets/MCRFileNodeServlet/DEHeft_derivate_00010214/2110100117004.pdf
- Steinhauer, H. W. (2017). *Samples, Weights, and Nonresponse: the Sample of Starting Cohort 3 of the National Educational Panel Study (Wave 7)*. Leibniz Institute for Educational Trajectories, National Educational Panel Study. Bamberg. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC3/7-0-1/SC3_7-0-1_W.7pdf
- Steinhauer, H. W. (2019). *Samples, Weights, and Nonresponse: the Sample of Starting Cohort 3 of the National Educational Panel Study (Wave 8)*. Leibniz Institute for Educational Trajectories, National Educational Panel Study. Bamberg. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC3/8-0-0/SC3_8-0-0_W8.pdf

Steinhauer, H. W., Aßmann, C., Zinn, S., Goßmann, S., & Rässler, S. (2015). Sampling and Weighting Cohort Samples in Institutional Contexts. *AStA Wirtschafts- und Sozialstatistisches Archiv*, 9(2), 131–157. <https://doi.org/10.1007/s11943-015-0162-0>

Steinhauer, H. W., & Zinn, S. (2016a). *NEPS Technical Report for Weighting: Weighting the Sample of Starting Cohort 3 of the National Educational Panel Study (Waves 1 to 3)* (Working Paper No. 63). Leibniz Institute for Educational Trajectories. Bamberg. https://www.neps-data.de/Portals/0/Working%20Papers/WP_LXIII.pdf

Steinhauer, H. W., & Zinn, S. (2016b). *NEPS Technical Report for Weighting: Weighting the Sample of Starting Cohort 3 of the National Educational Panel Study (Waves 1 to 5)* (Technical Report). Leibniz Institute for Educational Trajectories. Bamberg. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC3/5-0-0/SC3_5-0-0_W.pdf