

# Samples, Weights and Nonresponse: the Sample of Starting Cohort 3 of the National Educational Panel Study (Wave 6)

*Bahrani, S.; Steinhauer, H.W.*

Copyrighted Material  
Leibniz Institute for Educational Trajectories (LifBi)  
Wilhelmsplatz 3, 96047 Bamberg  
Director: Prof. Dr. Sandra Buchholz  
Executive Director of Research: Dr. Jutta von Maurice  
Executive Director of Administration: Dr. Robert Polgar  
Bamberg, 2017

# **Samples, Weights, and Nonresponse: the Sample of Starting Cohort 3 of the National Educational Panel Study (Wave 6)**

*Sara Bahrami, Leibniz Institute for Educational Trajectories  
Hans Walter Steinhauer, Leibniz Institute for Educational Trajectories*

Technical Report referring to [DOI:10.5157/NEPS:SC3:6.0.1](https://doi.org/10.5157/NEPS:SC3:6.0.1)

**E-mail address of lead author:**

methoden@lifbi.de

---

## 1 Prequel

This report complements NEPS Survey Paper No. 63 (Steinhauer & Zinn, 2016a) and gives details on Wave 6 of Starting Cohort 3 (SC3) of the National Educational Panel Study (NEPS). It refers to the Scientific Use File (SUF; DOI:10.5157/NEPS:SC3:6.0.1). 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 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)", giving the students status by the end of the wave. Finally, the last column gives 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, Aßmann, Zinn, Goßmann, and Rässler (2015). Details on calculating wave-specific nonresponse adjustments can be found in Steinhauer and Zinn (2016a) for waves 1 to 3 and Steinhauer and Zinn (2016b) for waves 1 to 5 respectively.

Table 1: Panel progress of SC3 by wave.

Wave (Time)	Study number	Sample	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	<b>Main</b>	6,112	0	6,112	5,778	334	0	13
2 (2011/2012)	A29, A57	<b>Main</b>	6,099	0	6,099	5,539	559	1	8
3 (2012/2013)	A30, A30A, A58	<b>All</b>	8,295	0	8,295	7,280	986	29	10
		Main	6,090	0	6,090	5,134	927	29	10
		Refr.	2,205	0	2,205	2,146	59	0	0
4 (2013/2014)	A31, A59	<b>All</b>	8,256	0	8,256	6,718	1,505	33	580
		Main	6,051	0	6,051	4,783	1,249	19	<sup>a</sup> 580
		Refr.	2,205	0	2,205	1,935	256	14	0
5 (2014/2015)	A94	<b>All</b>	7,643	0	7,643	5,778	1,625	240	0
		Main	5,452	0	5,452	4,001	1,273	178	0
		Refr.	2,191	0	2,191	1,777	352	62	0
6 (Spring 2015)	A98	<b>All</b>	7,403	0	7,403	5,586	1,740	77	0
		Main	5,274	0	5,274	3,920	1,293	61	0
		Refr.	2,129	0	2,129	1,666	447	16	0

Note.<sup>a</sup>: special-need students are excluded from the panel cohort after Wave 4

## 2 Changes in Wave 1 to Wave 5

In comparison to previous version some changes to the weights have been made. The calibrated weight ( $w_{t\_cal}$ ) for the main sample of SC3 is only for regular schools. For calibration Federal State, school type, and Gender have been used. Cross-sectional weights for Wave 1 and Wave 2 as well as longitudinal weights starting from Wave 1 are based on  $w_{t\_cal}$  instead of  $w_t$ . Weights for special-need students are still based on the nonresponse adjusted design weight ( $w_t$ ). For students of the main sample who have not dropped out of the sample together with the students of the refreshment sample a calibrated weight ( $w_{t3\_cal}$ ) is provided. For calibration Federal State and Gender have been used. From Wave 3 on cross-sectional weights and longitudinal weights starting from Wave 3 are based on  $w_{t3\_cal}$ . Finally, for all students who are not participating in a certain wave the corresponding weight is zero. The weights  $w_i$ ,  $w_t$ ,  $w_{t\_cal}$ , and  $w_{t3\_cal}$  are provided in a "pure" form, that is, they are neither trimmed nor standardized.

## 3 Participation in Wave 6

To account for the wave-specific participation decision of students response propensity reweighting is used to provide corresponding weights. To model binary participation decisions a model with probit link function is used, see Steinhauer and Zinn (2016a) and Steinhauer and Zinn (2016b) for details. The coefficients for the estimated models are displayed in Table 2. By Wave 6 the panel cohort has reduced to 7,326 students, see Table 1. As in Wave 5 some students have left their schools and entered vocational education. Like in Starting Cohort 4 of the NEPS these students (of both groups) are surveyed individually and are less likely to participate in Wave 6. Similar negative effects are found for students being in individual retracking. Furthermore having participated in previous waves significantly influences the participation decision in Wave 6 positively for both groups. For students being part of the refreshment sample, belonging to a school in federal states other than Berlin and Brandenburg have a negative influence on the participation decision.

## 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:SC3:6.0.1](https://doi.org/10.5157/NEPS:SC3:6.0.1). Besides individual/target ( $ID_t$ ) and institutional ( $ID_i$ ) identifiers, design information for the entire cohort is made available.<sup>1</sup> 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 "school type" ( $stratum\_imp1$ ), "Federal States" ( $stratum\_imp2\_R$ ), "regional classification" ( $stratum\_imp3\_R$ ) and "funding" ( $stratum\_imp4\_R$ ). In this release version additional information has been added to the design data, namely the total number of students ( $tx80113\_R$ ) and classes ( $tx80114\_R$ ) in Grade 8 in school year 2013/2014 as reported by official statistics. Nonresponse adjusted design weights on the institutional ( $w_i$ ) and the individual ( $w_t$ ) level are given for the entire cohort.<sup>2</sup> For all participants in a particular wave, cross-sectional weights are provided. Data

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

<sup>2</sup>The 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 2: Models estimating the individual participation propensities for students in Wave 6 of SC3 used to derive adjustment factors for adjusted wave-specific cross-sectional and longitudinal weights*

	Wave 6	
	Main sample	Refreshment sample
(Intercept)	−0.716*** (0.110)	−0.241 (0.269)
Explicit stratum Grade 5 but not Grade 7	−0.007 (0.079)	
Explicit stratum Remaining states		−0.259** (0.092)
Age group younger half	0.090 (0.049)	0.036 (0.070)
Native language Other than German		−0.102 (0.094)
Native language Missing		0.351 (0.458)
Gender female		−0.088 (0.069)
Student participated in Wave 2	0.200* (0.091)	
Student participated in Wave 3	0.390*** (0.076)	0.232 (0.225)
Student participated in Wave 4	0.630*** (0.068)	0.450*** (0.105)
Student participated in Wave 5	1.081*** (0.060)	1.120*** (0.091)
Student in individual retracking yes	−0.811*** (0.055)	−0.661*** (0.080)
Number of students	5,048	2,129

Notes: Reference categories are: Explicit stratum (SC3: Grade 5 and Grade 9), Age group (older half), Native language (German), Gender (male), Student participated in Wave t (no), Student in individual retracking (no). Students who have left the general school system prematurely and are also retracked by infas institute. To model individual participation, the `glm` function with a probit link provided in R (R Core Team, 2015) was used.

\*\*\*, \*\*, and \* denote significance at the 0.1%, 1%, and 5% level, respectively. Standard errors are given in parentheses.

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 raking.

*Table 3: Variables included in the weighting data of SC3 SUF version 6.0.1*

Variable	Applies to	Content
<i>Identifier</i>		
ID_t	8,317	Identifier for target person (students)
ID_i	8,317	Identifier for the institution (377 schools)
<i>Design information</i>		

Table 3: Variables included in the weighting data of SC3 SUF version 6.0.1

Variable	Applies to	Content
tstud_st	8,317	Study number the target person was first surveyed in (A28, A56, A63, A30A)
sample	8,317	Part of the sample the target person belongs to
stratum_exp	8,317	Explicit sampling stratum referring to the school
stratum_imp1	8,317	Implicit sampling stratum (school type according to sampling frame)
stratum_imp2_R	8,317	Implicit sampling stratum (Federal State the school is located in according to sampling frame)
stratum_imp3_R	8,317	Implicit sampling stratum (regional classification according to sampling frame)
stratum_imp4_R	8,317	Implicit sampling stratum (funding according to sampling frame)
tx80113_R	8,317	Total number of students in grade 8 as reported by official statistics
tx80114_R	8,317	Total number of classes in grade 8 as reported by official statistics
<i>Design weights adjusted for initial nonresponse</i>		
w_i	8,317	Weight for institution
w_t	8,317	Weight for target
w_t_cal	5,283	Weight for target, calibrated
w_t3_cal	8,054	Weight for targets in Wave 3, calibrated
<i>Weights adjusted for wave-specific nonresponse, standardized</i>		
w_t1	5,559	Cross-sectional weight for targets participating in Wave 1
w_t2	5,331	Cross-sectional weight for targets participating in Wave 2
w_t3	7,114	Cross-sectional weight for targets participating in Wave 3
w_t4	6,581	Cross-sectional weight for targets participating in Wave 4
w_t5	5,648	Cross-sectional weight for targets participating in Wave 5
w_t6	5,465	Cross-sectional weight for targets participating in Wave 6
w_t12	5,071	Longitudinal weight for targets participating in Wave 1 and 2
w_t123	4,516	Longitudinal weight for targets participating in Wave 1, 2, and 3
w_t1234	4,029	Longitudinal weight for targets participating in Wave 1, 2, 3, and 4
w_t12345	3,203	Longitudinal weight for targets participating in Wave 1, 2, 3, and 5
w_t123456	2,919	Longitudinal weight for targets participating in Wave 1, 2, 3, 5, and 6
w_t34	6,291	Longitudinal weight for targets participating in Wave 3 and 4
w_t345	5,119	Longitudinal weight for targets participating in Wave 3, 4, and 5



Table 3: Variables included in the weighting data of SC3 SUF version 6.0.1

Variable	Applies to	Content
w_t3456	4,601	Longitudinal weight for targets participating in Wave 3, 4, 5, and 6

**Acknowledgements** This paper uses data from the National Educational Panel Study (NEPS): Starting Cohort Grade 5, DOI:10.5157/NEPS:SC3:6.0.1. From 2008 to 2013, NEPS data was collected as part of the Framework Program for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, NEPS is carried out by the Leibniz Institute for Educational Trajectories (LifBi) at the University of Bamberg in cooperation with a nationwide network.

## References

- R Core Team. (2015). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from <https://www.R-project.org/>
- Statistisches Bundesamt. (2011). *Bildung und Kultur - Allgemeinbildende Schulen: Schuljahr 2010/2011* (Fachserie 11 No. Reihe 1). Wiesbaden: Statistisches Bundesamt. Retrieved 15.03.2017, from <https://www.destatis.de/DE/Publikationen/Thematisch/BildungForschungKultur/Schulen/AllgemeinbildendeSchulen2110100137004.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. doi: 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). Bamberg: Leibniz Institute for Educational Trajectories. Retrieved from [https://www.neps-data.de/Portals/0/Working%20Papers/WP\\_LXIII.pdf](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). Bamberg: Leibniz Institute for Educational Trajectories. Retrieved from [https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC3/5-0-0/SC3\\_5-0-0\\_W.pdf](https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC3/5-0-0/SC3_5-0-0_W.pdf)